

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 05:28:16 ; Search time 10901 Seconds

(without alignments)
11204.538 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818
Sequence: 1 agagaccatcatatga.....aaaaaaaaaaaaaaaa 2818

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seg length: 0

Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: *
1: gb_ba: *
2: gb_hvg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sta: *
12: gb_sy: *
13: gb_un: *
14: gb_vt: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sta: *
28: em_un: *
29: em_vt: *
30: em_hvg_hum: *
31: em_hvg_inv: *
32: em_hvg_other: *
33: em_hvg_mus: *
34: em_hvg_pin: *
35: em_hvg_rod: *
36: em_hvg_mam: *
37: em_hvg_vrt: *
38: em_sy: *
39: em_hvg_hum: *
40: em_hvg_mus: *
41: em_hvg_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2818	100.0	2818	6	BD016833
2	2791	99.0	2791	9	AB040431
3	2174.6	77.2	6564	6	BD016835
4	2174.6	77.2	11204	6	BD016860
5	2174.6	77.2	11204	9	AB040430
6	2174.6	77.2	71132	9	AC092184
7	2172	77.1	21132	6	BD016840
8	1818.6	64.5	1828	9	BC006296
9	603.4	21.4	2440	6	BD016828
10	603.4	21.4	2440	10	AF132979
11	595.4	21.1	597	12	AF529828
12	595	21.1	597	9	BT007402
13	595	21.1	597	12	BT008226
14	594.4	21.1	596	9	AF529815
15	594.4	21.1	596	9	AF529816
16	594.4	21.1	596	9	AF529819
17	594.4	21.1	596	9	AF529820
18	594.4	21.1	596	9	AF529821
19	594.4	21.1	596	9	AF529822
20	594.4	21.1	596	9	AF529823
21	594.4	21.1	596	9	AF529824
22	594.4	21.1	596	9	AF529825
23	594.4	21.1	596	9	AF529826
24	594.4	21.1	596	9	AF529827
25	594.4	21.1	596	12	AF529829
26	594.4	21.1	596	12	AF529830
27	594.4	21.1	596	12	AF529831
28	594.4	21.1	596	12	AF529833
29	594.4	21.1	596	12	AF529835
30	594.4	21.1	596	12	AF529837
31	594.4	21.1	596	12	AF529839
32	593.4	21.1	595	12	AF529834
33	592.8	21.0	596	9	AF529817
34	592.8	21.0	596	9	AF529818
35	592.8	21.0	596	12	AF529832
36	592.8	21.0	596	12	AF529836
37	592.8	21.0	596	12	AF529838
38	592.2	21.0	597	12	AF529841
39	592.2	21.0	597	12	AF529842
40	592.2	21.0	597	12	AF529843
41	592.2	21.0	597	12	AF529844
42	592.2	21.0	597	12	AF529845
43	592.2	21.0	597	12	AF529846
44	592.2	21.0	597	12	AF529847
45	592.2	21.0	597	12	AF529848

ALIGNMENTS

RESULT 1
LOCUS BD016833 2818 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel cytidine deaminase.
ACCESSION BD016833.1 GI:2258009
VERSION BD016833.1
KEYWORDS JP 2001245669-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2818)
AUTHORS Honjo, T. and Muramatsu, M.
TITLE Novel cytidine deaminase
JOURNAL Patent: JP 2001245669-A 6 11-SEP-2001;

COMMENT
JAPAN TOBACCO INC, TASUKU HONDO
OS Homo sapiens (human)
PN JP 2001245669-A/6
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981
PI TASUKU HONDO, MASAMICHI MURAMATSU
PC C12N15/09, A61K39/395, A61K39/395, A61P11/00, A61P11/06, A61P13/12,
PC A61P17/00,
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
C12N1/21
PC C12N5/10, C12N9/78, C12P21/02, C12P21/08// (C12N1/21, C12R1/19), PC
(C12N5/10, C12R1/91), C12N15/00, C12N5/00, (C12N5/00, C12R1/91) CC
FH Key Location/Qualifiers
FT 5'UTR (1) . . (79)
FT CDS (80) . . (676)
FT 3'UTR (677) . . (2818).
FEATURES
Source Location/Qualifiers
1. . 2818
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 2818; DB 6; Length 2818;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGAGACCATCATTAATTGAAGTGAATTTTCTGCTGAGACCTTGACGGAGGCAAGA 60
Db 1 AGGAGACCATCATTAATTGAAGTGAATTTTCTGCTGAGACCTTGACGGAGGCAAGA 60
QY 61 AGACACCTGGAGACCACTATGAGACAGCCTCTTGATGAACCGAGGAATTTCTTTACA 120
Db 61 AGACACCTGGAGACCACTATGAGACAGCCTCTTGATGAACCGAGGAATTTCTTTACA 120
QY 121 ATTCAAAATGTCCGCTGGGCTTAAGGATCGGCGTGAACCTTACTGTACTAGTGA 180
Db 121 ATTCAAAATGTCCGCTGGGCTTAAGGATCGGCGTGAACCTTACTGTACTAGTGA 180
QY 121 ATTCAAAATGTCCGCTGGGCTTAAGGATCGGCGTGAACCTTACTGTACTAGTGA 180
Db 121 ATTCAAAATGTCCGCTGGGCTTAAGGATCGGCGTGAACCTTACTGTACTAGTGA 180
QY 181 GAGGCGTGACAGTGTACATCTTTTCACTGACCTTTGTTATCTTCGATAGAACG 240
Db 181 GAGGCGTGACAGTGTACATCTTTTCACTGACCTTTGTTATCTTCGATAGAACG 240
QY 181 GAGGCGTGACAGTGTACATCTTTTCACTGACCTTTGTTATCTTCGATAGAACG 240
Db 181 GAGGCGTGACAGTGTACATCTTTTCACTGACCTTTGTTATCTTCGATAGAACG 240
QY 241 CTGCGCAGTGAATGTCTCTCTCCGCTACATCTCGACCTGACCTGACCTGCGCG 300
Db 241 CTGCGCAGTGAATGTCTCTCTCCGCTACATCTCGACCTGACCTGACCTGCGCG 300
QY 301 CTGCTACCGGCTCACCGTGTACCTCTGAGAGCCCTGCTACAGCTGTGCCGACATGT 360
Db 301 CTGCTACCGGCTCACCGTGTACCTCTGAGAGCCCTGCTACAGCTGTGCCGACATGT 360
QY 361 GGGCGACCTTTCTGCGAGGGAACCCCAACTCAGTCTGAGAGCTTTCACCGCGGCTCTA 420
Db 361 GGGCGACCTTTCTGCGAGGGAACCCCAACTCAGTCTGAGAGCTTTCACCGCGGCTCTA 420
QY 421 CTCTCTGAGAGACCGCAAGGCTGAGCCGAGGGGCTGCGGCGGCTGCAACCGCGGGGT 480
Db 421 CTCTCTGAGAGACCGCAAGGCTGAGCCGAGGGGCTGCGGCGGCTGCAACCGCGGGGT 480
QY 481 GCAATATAGCATCATGACCTTCAAGATTATTTTACTGCTGGAATCTTTTGTGAAAA 540
Db 481 GCAATATAGCATCATGACCTTCAAGATTATTTTACTGCTGGAATCTTTTGTGAAAA 540
QY 541 CCATGAAAGACCTTCAAGAGCTGAGAGGGCTGACATGAAATTCAGTCTCTCCAG 600
Db 541 CCATGAAAGACCTTCAAGAGCTGAGAGGGCTGACATGAAATTCAGTCTCTCCAG 600
QY 601 ACAAGCTTGGCGGACCTTTTGGCCCTGTATGAGGTTGATGACTTACGAGACGATTTG 660
Db 601 ACAAGCTTGGCGGACCTTTTGGCCCTGTATGAGGTTGATGACTTACGAGACGATTTG 660
QY 661 TACTTTGGGACTTGTATGACACTTCCAGGAATGTACACAGAGAAATATCTCTGCTG 720

Db 661 TACTTTGGGACTTGTATGACACTTCCAGGAATGTACACAGAGAAATATCTCTGCTG 720
QY 721 AAGACATGATATAAAACAGCTCTTCAATCTTCTCTGTTTATTTCTTCAACTCTCAC 780
Db 721 AAGACATGATATAAAACAGCTCTTCAATCTTCTCTGTTTATTTCTTCAACTCTCAC 780
QY 781 TTTCTTGAAGTTTACAGAAAAATTTATATAGACTCTTTAAAGATCTATGCTCTG 840
Db 781 TTTCTTGAAGTTTACAGAAAAATTTATATAGACTCTTTAAAGATCTATGCTCTG 840
QY 841 AAAATAGAGAGAGAACACAGGCTGCGCAGAGAGCTGCTCAATTTGCTGAGTTTGAAT 900
Db 841 AAAATAGAGAGAGAACACAGGCTGCGCAGAGAGCTGCTCAATTTGCTGAGTTTGAAT 900
QY 901 GCAACATTTGCTCCCTATCTGGGAATTAACAGAACTGACGACCTTGAAGCATCTTAAAGTGT 960
Db 901 GCAACATTTGCTCCCTATCTGGGAATTAACAGAACTGACGACCTTGAAGCATCTTAAAGTGT 960
QY 961 CAACGTTTTCTATGACTTTTAGGTAGAGAGAGAGAGATGATCCCTAAAGCATG 1020
Db 961 CAACGTTTTCTATGACTTTTAGGTAGAGAGAGAGAGATGATCCCTAAAGCATG 1020
QY 1021 GTGAGAGATCAAAATGTTTATATCAACATCTTATTTATTTGATTTCAATTTGATTAAC 1080
Db 1021 GTGAGAGATCAAAATGTTTATATCAACATCTTATTTATTTGATTTCAATTTGATTAAC 1080
QY 1081 AGTGTGTATGATATGATTTTCTATTTCTTTCCCTTGAAGTTTCAATTTCAATTAAC 1140
Db 1081 AGTGTGTATGATATGATTTTCTATTTCTTTCCCTTGAAGTTTCAATTTCAATTAAC 1140
QY 1141 AAATCTTCCATCAGGCGATGATCTATAGAGCTCTCTAATGAGATATCTGGGTATGT 1200
Db 1141 AAATCTTCCATCAGGCGATGATCTATAGAGCTCTCTAATGAGATATCTGGGTATGT 1200
QY 1201 GACCCCAACATCTCTCCCAAGCATTAATCAATCAATGCGCTGATTTTATATAC 1260
Db 1201 GACCCCAACATCTCTCCCAAGCATTAATCAATCAATGCGCTGATTTTATATAC 1260
QY 1261 CAGAGCATGTTTATATGTTTATCAAAAGAAATGTTATGGGTGGAGATGAGATATAT 1320
Db 1261 CAGAGCATGTTTATATGTTTATCAAAAGAAATGTTATGGGTGGAGATGAGATATAT 1320
QY 1321 GACCATCATGCTACCTTCAAGCTTCTTAAATGAAGATCTTAAATGGGCGAGAGAC 1380
Db 1321 GACCATCATGCTACCTTCAAGCTTCTTAAATGAAGATCTTAAATGGGCGAGAGAC 1380
QY 1381 TGTGAACAAGACACCCCTAATATGAGTGTATGCTGAGAGAAATCTTCTGGAACCG 1440
Db 1381 TGTGAACAAGACACCCCTAATATGAGTGTATGCTGAGAGAAATCTTCTGGAACCG 1440
QY 1441 AAATCTTTTAAAGAAAGTCCCTAATTTAGAAACACCAAACTTCAATATCATATTA 1500
Db 1441 AAATCTTTTAAAGAAAGTCCCTAATTTAGAAACACCAAACTTCAATATCATATTA 1500
QY 1501 GCAAAATTTGAAGAAAGTCTTGAATGTTGGGAGAGAAATCTATGGCTCTGCT 1560
Db 1501 GCAAAATTTGAAGAAAGTCTTGAATGTTGGGAGAGAAATCTATGGCTCTGCT 1560
QY 1561 GGGTCTTTATCTCAAGAAATGCAATCAAGTTCAGATTTTGTATGTGTGT 1620
Db 1561 GGGTCTTTATCTCAAGAAATGCAATCAAGTTCAGATTTTGTATGTGTGT 1620
QY 1621 GATGCTTCTCCCAAGATATATTAATATATAGAGTTTGAACAAACAGAAATGATTA 1680
Db 1621 GATGCTTCTCCCAAGATATATTAATATATAGAGTTTGAACAAACAGAAATGATTA 1680
QY 1681 AGCTGCAACCGTGGACACAGCTCATATGTTCTAGCTCTTGGAGGTTGAGAGGAGGA 1740
Db 1681 AGCTGCAACCGTGGACACAGCTCATATGTTCTAGCTCTTGGAGGTTGAGAGGAGGA 1740
QY 1741 TGGCTTGAACAAGGTTTCAAGGCAAGCTGGGCAACATTAACAGATCTGTCTTCAA 1800
Db 1741 TGGCTTGAACAAGGTTTCAAGGCAAGCTGGGCAACATTAACAGATCTGTCTTCAA 1800

```

QY 1801 AAAAAAAAAAAAAAAAAAGAGAGAGGCGCGCGTGTGGCTCAAGCCTGTATATCC 1860
Db 1801 AAAAAAAAAAAAAAAAAAGAGAGAGGCGCGCGTGTGGCTCAAGCCTGTATATCC 1860
QY 1861 CAGACCTTTGGAGCGCGGCGGATCACCCTGTGTGTCAGAGTTTGAACCGCCT 1920
Db 1861 CAGACCTTTGGAGCGCGGCGGATCACCCTGTGTGTCAGAGTTTGAACCGCCT 1920
QY 1921 GGGCAATGAGCAAAACCCGCTGTATCAAAATGCAAAATTTAGCAGGCGGTGTAC 1980
Db 1921 GGGCAATGAGCAAAACCCGCTGTATCAAAATGCAAAATTTAGCAGGCGGTGTAC 1980
QY 1981 AGGCACTGTATATCCAGCTACTTGGAGGCTGAGGAGAGATGCTTGAACCGAGA 2040
Db 1981 AGGCACTGTATATCCAGCTACTTGGAGGCTGAGGAGAGATGCTTGAACCGAGA 2040
QY 2041 GGTGAGGTTGCAATGAGCTGAGATGCTGCGCTTGCATCTCAGCCTGGGCGACAGAGA 2100
Db 2041 GGTGAGGTTGCAATGAGCTGAGATGCTGCGCTTGCATCTCAGCCTGGGCGACAGAGA 2100
QY 2101 AGACTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
Db 2101 AGACTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
QY 2161 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
Db 2161 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
QY 2221 GCCAATGAGGAGATCCCTATTTGTCTTTTGTGTATTTGTCTTCAAACTGCTTT 2280
Db 2221 GCCAATGAGGAGATCCCTATTTGTCTTTTGTGTATTTGTCTTCAAACTGCTTT 2280
QY 2281 GACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
Db 2281 GACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
QY 2341 TGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
Db 2341 TGCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
QY 2401 TTATGTACATAGTTTGTAAAAAGAGTTTGTATTTGTATTTGTATTTGTATTTAT 2460
Db 2401 TTATGTACATAGTTTGTAAAAAGAGTTTGTATTTGTATTTGTATTTGTATTTAT 2460
QY 2461 TTATATTTATTTTGGCTGTATATTTTATTTATTTATTTATTTATTTATTTATTTGA 2520
Db 2461 TTATATTTATTTTGGCTGTATATTTTATTTATTTATTTATTTATTTATTTATTTGA 2520
QY 2521 AATGAGGCTCAAGGTTCAATAATTTTATTAATTTTATTAATTTTATTAATTTTATTA 2580
Db 2521 AATGAGGCTCAAGGTTCAATAATTTTATTAATTTTATTAATTTTATTAATTTTATTA 2580
QY 2581 GTAATGTGAATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2640
Db 2581 GTAATGTGAATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2640
QY 2641 TTTATGACAGCAAAATTTGCTTGGCTCACTTTCAATCAGTTAAATTAATTAATTAAT 2700
Db 2641 TTTATGACAGCAAAATTTGCTTGGCTCACTTTCAATCAGTTAAATTAATTAATTAAT 2700
QY 2701 TTTGAGAGCTGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2760
Db 2701 TTTGAGAGCTGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2760
QY 2761 TAAAAATAAAAATTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2818
Db 2761 TAAAAATAAAAATTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2818

```

RESULT 2
AB040431
LOCUS AB040431 2791 bp mRNA linear PRI 03-OCT-2000

```

DEFINITION Homo sapiens AID mRNA for activation-induced cytidine deaminase,
complete CDS.
ACCESSION AB040431
VERSION AB040431.1 GI:9988409
KEYWORDS AID; activation-induced cytidine deaminase; Human AID.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (sites)
Muto,T., Muramatsu,M., Taniwaki,M., Kinoshita,K. and Honjo,T.
Isolation, tissue distribution, and chromosomal localization of the
human activation-induced cytidine deaminase (Aid) gene
Genomics 68 (1), 85-88 (2000)
MEDLINE 20408890
PUBMED 10950930
REFERENCE 2 (sites)
Revy,P., Muto,T., Levy,Y., Geisemann,F., Plebani,A., Sanal,O.,
Catalan,N., Forveille,M., Dufourcq-Lagelouse,R., Gennery,A.,
Tzscan,I., Eszy,F., Kayserili,H., Ugazio,A.G., Brousse,N.,
Muramatsu,M., Notarangelo,L.D., Kinoshita,K., Honjo,T., Fischer,A.
and Durandy,A.
Activation-induced cytidine deaminase (Aid) deficiency causes the
autosomal recessive form of the Hyper-Igm syndrome (HIGM2)
Cell 102 (5), 565-575 (2000)
JOURNAL MEDLINE 20460541
PUBMED 11007475
TITLE 3 (bases 1 to 2791)
Muto,T., Muramatsu,M., Taniwaki,M., Kinoshita,K. and Honjo,T.
Direct Submission
Submitted (18-MAR-2000) Tasuku Honjo, Kyoto University, Department
of Medical Chemistry, Faculty of Medicine, Yoshida, Sakyo-ku,
Kyoto, Kyoto 606-8501, Japan (E-mail: honjo@med.kyoto-u.ac.jp,
Tel:81-75-753-4371 (ex.4371), Fax:81-75-753-4388)
FEATURES
location/Qualifiers
1..2791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2791
/gene="AID"
77..673
/gene="AID"
/codon_start=1
/product="activation-induced cytidine deaminase"
/protein_id="BAI2721.1"
/db_xref="GI:9988410"
/translation="MPSLMNRKRPKYGRNVRMAKGRRETYCTVYKRDATRSFL
DFGLRKNQCHVELFLNRTISDMDLPKRCRVVFTSMPQYDCAHVAADLRNP
NLSLRIFTALVYCEDRKAEPEGLRIHAGVOIALMTFKDYFCWNTVEVNERFEK
AMEGLHENSVRSLRQLRILLPLYEVDDLRDAFRLTGL"
ORIGIN
Query Match 99.0%; Score 2791; DB 9; Length 2791;
Bseq Local Similarity 100.0%; Pred.No. 0;
Matches 2791; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 GAACCATCATTAATGAGAGATTTTCTGGCTGAGCTTGAGGAGGAGAGAGAGAGAGAG 63
Db 1 GAACCATCATTAATGAGAGATTTTCTGGCTGAGCTTGAGGAGGAGAGAGAGAGAGAG 60
QY 64 CACTGTGACACCACTATGAGACAGCTCTTGATGAACCGAGAGAGATTTCTTTACCAATT 123
Db 61 CACTGTGACACCACTATGAGACAGCTCTTGATGAACCGAGAGAGATTTCTTTACCAATT 120
QY 124 CAATAATGTCGCTGGGCTTAAGGCTGGCGTGAACCTTCTGTCTACGTAGTGAAGAG 183
Db 121 CAATAATGTCGCTGGGCTTAAGGCTGGCGTGAACCTTCTGTCTACGTAGTGAAGAG 180
QY 184 GCGTGAAGTGCATACCTCTTTCACTGACATTTGTTATCTTGGCAATTAAGACGGCTG 243
Db 181 GCGTGAAGTGCATACCTCTTTCACTGACATTTGTTATCTTGGCAATTAAGACGGCTG 240

```

QY 244 CCAAGTGAATGTCTCTTCCGCTACATCTGGACCTGAGACCTTGCGCTG 303
Db 241 CCAAGTGAATGTCTCTTCCGCTACATCTGGACCTGAGACCTTGCGCTG 300
QY 304 CTACCGGTCACTGTGTACCTCTGAGACCTCTGCTACATGCTGCGCATGTGGC 363
Db 301 CTACCGGTCACTGTGTACCTCTGAGACCTCTGCTACATGCTGCGCATGTGGC 360
QY 364 CGACTTCTGCGAGGAGAACCCCACTCAGTCTGAGATCTTACCGCGCGCTCACT 423
Db 361 CGACTTCTGCGAGGAGAACCCCACTCAGTCTGAGATCTTACCGCGCGCTCACT 420
QY 424 CTGTGAGACCGCAAGGCTGAGCCGAGGGCTGCGCGCTGCAACCGCGCGGTGCA 483
Db 421 CTGTGAGACCGCAAGGCTGAGCCGAGGGCTGCGCGCTGCAACCGCGCGGTGCA 480
QY 484 AATAGCCATCTAGACCTTCAAAAGATATTTTACTGCTGGAATCTTTGTAGAAAACA 543
Db 481 AATAGCCATCTAGACCTTCAAAAGATATTTTACTGCTGGAATCTTTGTAGAAAACA 540
QY 544 TGAAGAAGCTTCAAGCTGGGAGAGGCTGCAATGAAATTTCACTGCTCTCCAGACA 603
Db 541 TGAAGAAGCTTCAAGCTGGGAGAGGCTGCAATGAAATTTCACTGCTCTCCAGACA 600
QY 604 GCTTCGGGCACTCTTTTGGCCCTGTATGAGGTGATGACTTACGAGACGCAATTTGCTAC 663
Db 601 GCTTCGGGCACTCTTTTGGCCCTGTATGAGGTGATGACTTACGAGACGCAATTTGCTAC 660
QY 664 TTTGGAGCTTTGATAGCACTTCCAGGAATGTCAACACGATGAAATATCTCTGTGAAG 723
Db 661 TTTGGAGCTTTGATAGCACTTCCAGGAATGTCAACACGATGAAATATCTCTGTGAAG 720
QY 724 ACATGTGATAAAAACAGTCTTCAAGTCTTCTGTGTTTATTTCTTCAACTCTCACTTT 783
Db 721 ACATGTGATAAAAACAGTCTTCAAGTCTTCTGTGTTTATTTCTTCAACTCTCACTTT 780
QY 784 CTTAGAGTTTACAGAAAAATTTATATAGACTCTTAAAGAACTATGCTGAGAA 843
Db 781 CTTAGAGTTTACAGAAAAATTTATATAGACTCTTAAAGAACTATGCTGAGAA 840
QY 844 ATAGAGAGAGAACACAGGCTGCGCAGAGAGCTGCTGCAATGCTGCAATTTGAATGCA 903
Db 841 ATAGAGAGAGAACACAGGCTGCGCAGAGAGCTGCTGCAATGCTGCAATTTGAATGCA 900
QY 904 ACATGTCTCTTCTGAGAAATACAGAACTGCAAGACCTTGGAGCATCTTAAAGTGTCA 963
Db 901 ACATGTCTCTTCTGAGAAATACAGAACTGCAAGACCTTGGAGCATCTTAAAGTGTCA 960
QY 964 CGTTTTCATAGACTTTTATAGTATGAGAGAGAGAGTATCTTAAAGAACTATGCTGAG 1023
Db 961 CGTTTTCATAGACTTTTATAGTATGAGAGAGAGAGTATCTTAAAGAACTATGCTGAG 1020
QY 1024 AGAGAGATCAATGTTTATATCAACATCTTATTTATTTGATTTCAATTTGATTAACAGT 1083
Db 1021 AGAGAGATCAATGTTTATATCAACATCTTATTTATTTGATTTCAATTTGATTAACAGT 1080
QY 1084 GGTGTATAGATATGATTTTCTATTTCTTCCCTTGAAGTTTCACTTCAAGTAACAA 1143
Db 1081 GGTGTATAGATATGATTTTCTATTTCTTCCCTTGAAGTTTCACTTCAAGTAACAA 1140
QY 1144 CTCTTCCATCAGGCGATGATCTATAGACCTCTCAATGAGAGTATCTGGGTGATGTGAC 1203
Db 1141 CTCTTCCATCAGGCGATGATCTATAGACCTCTCAATGAGAGTATCTGGGTGATGTGAC 1200
QY 1204 CCCAAGCATCTCTCCAAAGATTAATATCAATCATGCGCTATGTTTATCTACGAG 1263
Db 1201 CCCAAGCATCTCTCCAAAGATTAATATCAATCATGCGCTATGTTTATCTACGAG 1260
QY 1264 AAGCATGTTTATGTTGTACAAAAGAGATTTATGAGGTGGAGTGGAGTATAGAC 1323
Db 1261 AAGCATGTTTATGTTGTACAAAAGAGATTTATGAGGTGGAGTGGAGTATAGAC 1320
QY 1324 CATGATGTGACCTTCAAGCTACTTTAATTAAGATCTTAAATGGCGAGAGACTGT 1383

Db 1321 CATGATGTGACCTTCAAGCTACTTTAATTAAGATCTTAAATGGCGAGAGACTGT 1380
QY 1384 GAAACAAGACACCTTATATATGAGTTGATGTCTGAAGTACAAATCTTCTGAAAACGCAAA 1443
Db 1381 GAAACAAGACACCTTATATATGAGTTGATGTCTGAAGTACAAATCTTCTGAAAACGCAAA 1440
QY 1444 CTCTTTAAGAGAGCCCTAATTTAGAAACCCCAAACTTGACATATCAATTAATGCA 1503
Db 1441 CTCTTTAAGAGAGCCCTAATTTAGAAACCCCAAACTTGACATATCAATTAATGCA 1500
QY 1504 AACATTTGAAGAGAGTTGCTTGAATGTGGAGAGAGAAATCTATTTGCTCTGAGG 1563
Db 1501 AACATTTGAAGAGAGTTGCTTGAATGTGGAGAGAGAGAAATCTATTTGCTCTGAGG 1560
QY 1564 TCTCTTCACTCAGAAATGCCAATCAGGTCAAGTTTGTCAATTTGTATGTGTGAT 1623
Db 1561 TCTCTTCACTCAGAAATGCCAATCAGGTCAAGTTTGTCAATTTGTATGTGTGAT 1620
QY 1624 GCTTCTCCGAAGATATTTAATCTATTAAGAGTTGTGACAAACAGAAATGATAAGC 1683
Db 1621 GCTTCTCCGAAGATATTTAATCTATTAAGAGTTGTGACAAACAGAAATGATAAGC 1680
QY 1684 TGGCAACCGTGGCAACAGCTCATAGTTCTAGCTCTTGGAGGTTGAGAGAGAGATGG 1743
Db 1681 TGGCAACCGTGGCAACAGCTCATAGTTCTAGCTCTTGGAGGTTGAGAGAGAGATGG 1740
QY 1744 CTGAAACAGGTGTTCAGAGGCAAGCTGAGCAATTAACAGATCTGTCTCAAAAA 1803
Db 1741 CTGAAACAGGTGTTCAGAGGCAAGCTGAGCAATTAACAGATCTGTCTCAAAAA 1800
QY 1804 AAAAAAAAAAAAAAGAAAGAGAGAGGCGCGGTGTGCTACGCTGTAACTCCAG 1863
Db 1801 AAAAAAAAAAAAAAGAAAGAGAGAGGCGCGGTGTGCTACGCTGTAACTCCAG 1860
QY 1864 CACTTGGAGAGCGGAGCGGAGCTGATCACTGTGTGAGAGTTGAGACCGAGCTGGC 1923
Db 1861 CACTTGGAGAGCGGAGCGGAGCTGATCACTGTGTGAGAGTTGAGACCGAGCTGGC 1920
QY 1924 CAACATGCAAAACCCGCTGTATCTCAAAATGCAAAATTTAGCAGGCTGTGAGCAG 1983
Db 1921 CAACATGCAAAACCCGCTGTATCTCAAAATGCAAAATTTAGCAGGCTGTGAGCAG 1980
QY 1984 CACTGTATCTCAGCTTCTTGGAGGCTGAGGCAAGAGATGCTTGAACCCAGAGGT 2043
Db 1981 CACTGTATCTCAGCTTCTTGGAGGCTGAGGCAAGAGATGCTTGAACCCAGAGGT 2040
QY 2044 GAGGTTGCAATAGCTGAGATCTGTCCGTTTGCATCTCAGCTGGGCGACAAGGCAAGA 2103
Db 2041 GAGGTTGCAATAGCTGAGATCTGTCCGTTTGCATCTCAGCTGGGCGACAAGGCAAGA 2100
QY 2104 CTCTGTCTCAGAAAAAAG 2163
Db 2101 CTCTGTCTCAGAAAAAAG 2160
QY 2164 GAGAGAGTGGGAGAGATTCAGAGAAATTTGCTTATCAACAAATGTAAAGAGCC 2223
Db 2161 GAGAGAGTGGGAGAGATTCAGAGAAATTTGCTTATCAACAAATGTAAAGAGCC 2220
QY 2224 AATAAGGATCTCCTATTTGTCTCTTTGTGTCTATTTGTCCCTAACAACTGTCTTGAC 2283
Db 2221 AATAAGGATCTCCTATTTGTCTCTTTGTGTCTATTTGTCCCTAACAACTGTCTTGAC 2280
QY 2284 AGTGAAGAAATATTCAGAAATCAATATCCCTGTGCGCTTATTAAGCAACCTTGC 2343
Db 2281 AGTGAAGAAATATTCAGAAATCAATATCCCTGTGCGCTTATTAAGCAACCTTGC 2340
QY 2344 AATGAAGTGAAGCATCCACAGAAACTTGAATGACAACTGTCTTATTTATCTTA 2403
Db 2341 AATGAAGTGAAGCATCCACAGAAACTTGAATGACAACTGTCTTATTTATCTTA 2400
QY 2404 TTGTACATAGTTTGTAAAGAGTTAAATTTGTACTTCAATGATTTATTTTAA 2463

Db	2401	TTGTACATTAAGTTTGTAAAGGTTAAAGTTACTTCATGTATTCATTATATTTTA	2460		
QY	2464	TATTAATTTTGGCGTGAATGATTTTTTTTATTAACATGATTTCCTTTCGTATATATGAAT	25223		
Db	2461	TATTAATTTTGGCGTGAATGATTTTTTTTATTAACATGATTTCCTTTCGTATATATGAAT	25202		
QY	2524	GGAGCTCGAAGCTTCATTAATTTTAACTTTAGAAAGATTTCTAATTAACAACGTATGTA	25833		
Db	2521	GGAGCTCGAAGCTTCATTAATTTTAACTTTAGAAAGATTTCTAATTAACAACGTATGTA	25803		
QY	2584	ATTGTAAACATGTCAGTAAATGGTGCACGAGCAATTCCTCTGATTTTTTTAGTAAACCTTT	26443		
Db	2581	ATTGTAAACATGTCAGTAAATGGTGCACGAGCAATTCCTCTGATTTTTTTAGTAAACCTTT	26403		
QY	2644	ATGACAGCAAAATTTTGCTTCCTGGCTCACTTCACATTCAGTTAAATAAATGATAATATTTT	27030		
Db	2641	ATGACAGCAAAATTTTGCTTCCTGGCTCACTTCACATTCAGTTAAATAAATGATAATATTTT	27000		
QY	2704	GGAACCTGTGAAGATTAATAATACCAATTAATTAATTAAGTATTAATGAAGTTAA	27633		
Db	2701	GGAACCTGTGAAGATTAATAATACCAATTAATTAATTAAGTATTAATGAAGTTAA	27603		
QY	2764	AAAAAAAAATCAGTATGATGGAATTAACCTGG	2794		
Db	2761	AAAAAAAAATCAGTATGATGGAATTAACCTGG	2791		
RESULT 3	BD016835	6564 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD016835				
DEFINITION	Novel cytidine deaminase.				
ACCESSION	BD016835				
VERSION	BD016835.1				
KEYWORDS	JP 2001245669-A/8.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 6564)				
JOURNAL	Hongo, T. and Muramatsu, M.				
COMMENT	Novel cytidine deaminase				
	Patent: JP 2001245669-A 8 11-SEP-2001;				
	USPAT TOBACO INC, TASUKU HONDO				
	OS Homo sapiens (human)				
	PN JP 2001245669-A/8				
	PD 11-SEP-2001				
	PF 28-MAR-2000 JP 2000092981				
	PI TASUKU HONDO, KASAMICHI MURAMATSU				
	PC C12N15/00, A61K39/395, A61P1/00, A61P11/06, A61P13/12,				
	PC A61P17/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC				
	C12N1/21,				
	PC C12N5/10, C12N9/78, C12P21/02, C12P21/08/(C12N1/21, C12R1:19), PC				
	(C12N5/10, C12R1:91), C12N15/00, C12N5/00, (C12N5/00, C12R1:91) CC				
FEATURES	PH	Key	Location/Qualifiers.		
source	1..6564	Location/Qualifiers			
ORIGIN	1..6564	Location/Qualifiers			
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
Query Match	77.2%	Score 2174.6;	DB 6;	Length 6564;	
Best Local Similarity	99.8%;	Pred. No. 0;			
Matches 2177;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;	
QY	617	CTTTTGGCCCTGTATGAGTTGATGACTTACGAGACGCAATTCGTACTTTGGGACTTTGA	676		
Db	3735	CTACAGCCCCGTATGAGTTGATGACTTACGAGACGCAATTCGTACTTTGGGACTTTGA	3794		
QY	677	TAGCAACTTCAGGAATGTCAACAACGATGAATATCTCTGCTGAAGACAGTGGATTA	736		
Db	3795	TAGCAACTTCAGGAATGTCAACAACGATGAATATCTCTGCTGAAGACAGTGGATTA	3854		

QY	73	TAACAGTCCCTCAAGCTCTCTGTTTTTATCTTCAACTCCACTCTTCTTAGAGTTTACA	798
Db	3855	AACAGTCTTCAAGCTCTTCTGTTTTATCTTCAACTCCACTCTTCTTAGAGTTTACA	391
QY	797	GAATAAATAATTAATATACGACTCTTTAAAGAATCATGTATGCTTTGAAATAGAGAAGAAC	856
Db	3915	GAATAAATAATTAATATACGACTCTTTAAAGAATCATGTATGCTTTGAAATAGAGAAGAAC	397
QY	85	ACAGGCTGGCCAGGACGTGCTGCATTTGGTCAGTTTGAATGCAACATTTGCCCTTA	916
Db	3975	ACAGGCTGGCCAGGACGTGCTGCATTTGGTCAGTTTGAATGCAACATTTGCCCTTA	403
QY	917	CTGGGAATTAACAAGCATGACAGGACCTGGGAGAGATCCTAAAGTGTACAAGTTTCTATAGA	976
Db	4035	CTGGGAATTAACAAGCATGACAGGACCTGGGAGAGATCCTAAAGTGTACAAGTTTCTATAGA	409
QY	977	CTTTTAGGTAGATGAGACAGAAAGTATGATCTTTAAAGAATGAGTGTAGAGAGATCAAAATG	103
Db	4095	CTTTTAGGTAGATGAGACAGAAAGTATGATCTTTAAAGAATGAGTGTAGAGAGATCAAAATG	415
QY	1037	TTTTTAATACAACATCCTTTATTTATTTGATTCATTGATTAACAAGTGTGTATGATATA	1096
Db	4155	TTTTTAATACAACATCCTTTATTTATTTGATTCATTGATTAACAAGTGTGTATGATATA	421
QY	1097	GATTTTCTATCTTTTCCCTTGAAGTTTACTTTCAATTAACAACAACTCTTCCATGAC	1155
Db	4215	GATTTTCTATCTTTTCCCTTGAAGTTTACTTTCAATTAACAACAACTCTTCCATGAC	427
QY	1157	CCATGATCTATAGACCTCTTAATAGAGATCTGGGTGATTTGTAACCCCAACCATATGC	121
Db	4275	CCATGATCTATAGACCTCTCTTAATAGAGATCTGGGTGATTTGTAACCCCAACCATATGC	433
QY	1217	TCCAAAGCATTAATATCCAAATCATGCGCTGATGTTTTAATCAGAGAGAAGATGTTTTTA	1276
Db	4335	TCCAAAGCATTAATATCCAAATCATGCGCTGATGTTTTAATCAGAGAGAAGATGTTTTTA	439
QY	1277	TGTTTGTACAAAAGAAAGTTTGTATGGGTGGGGATGAGAGTATAGACATGATGATGATAC	1338
Db	4395	TGTTTGTACAAAAGAAAGTTTGTATGGGTGGGGATGAGAGTATAGACATGATGATGATAC	445
QY	1337	CTTCAAGCTACTTTAATTAAGAAGATCTTTAAATGAGGACAGAGACCTGTGAACAAGACACC	1396
Db	4455	CTTCAAGCTACTTTAATTAAGAAGATCTTTAAATGAGGACAGAGACCTGTGAACAAGACACC	451
QY	1397	TAAATATGGTGTGATGTCGTGAAGTACAAATCTTGTGAAACGCAAACTCTTTTAAGGAA	1456
Db	4515	TAAATATGGTGTGATGTCGTGAAGTACAAATCTTGTGAAACGCAAACTCTTTTAAGGAA	457
QY	1457	GTCCTTAATTTAGAAACACCCCAAACTTCCATATCAATTAATTAAGCAACATTTGAAGG	1516
Db	4575	GTCCTTAATTTAGAAACACCCCAAACTTCCATATCAATTAATTAAGCAACATTTGAAGG	463
QY	1517	AAGTTGCTGGAATGTGGGAGAGAGAAATCTATTTGGCTCGAGGGCTCTTCAATCTCA	1576
Db	4635	AAGTTGCTGGAATGTGGGAGAGAGAAATCTATTTGGCTCGAGGGCTCTTCAATCTCA	469
QY	1577	GAATATGCCAATCAGGTCAGAGTTTGCTACATTTTGTATGTGTGTGATGCTTCTCCAAAG	1636
Db	4695	GAATATGCCAATCAGGTCAGAGTTTGCTACATTTTGTATGTGTGTGATGCTTCTCCAAAG	475
QY	1637	GTAATATTAACATATTAAGAGATGTGTGACAAAACAGAAATGATTAAGCTGCGAACCCGTGGC	1696
Db	4755	GTAATATTAACATATTAAGAGATGTGTGACAAAACAGAAATGATTAAGCTGCGAACCCGTGGC	481
QY	1697	ACAGGCTCATATGTTCTAGTGTCTGGAGAGTGTGAGAGGAGATGCTTGAACACAGGT	1756
Db	4815	ACAGGCTCATATGTTCTAGTGTCTGGAGAGTGTGAGAGGAGATGCTTGAACACAGGT	487
QY	1757	GTTCAAGGCCACGCTGGGCAACATTAACAAGATCTGTCTCTCAAAAAAAAAAAAAAAAAA	1816
Db	4875	GTTCAAGGCCACGCTGGGCAACATTAACAAGATCTGTCTCTCAAAAAAAAAAAAAAAAAA	493

QY 1817 AAGAAAGAGAGAGGCGGCGGTGGTCAAGCCCTGTAATCCAGACCTTTGGAGGC 1876
Db 4935 AAGAAAGAGAGAGGCGGCGGTGGTCAAGCCCTGTAATCCAGACCTTTGGAGGC 4994
QY 1877 CGAGCCGCGGAGATCACCCTGTGGTCAAGAGTTGAGCCAGCTGGCCAAATGGCAAA 1936
Db 4995 CGAGCCGCGGAGATCACCCTGTGGTCAAGAGTTGAGCCAGCTGGCCAAATGGCAAA 5054
QY 1937 CCGCGTGTACTCAAAATGCAAAAATTAGCCAGCGGTGGTACAGGCACTGTATGCC 1996
Db 5055 CCGCGTGTACTCAAAATGCAAAAATTAGCCAGCGGTGGTACAGGCACTGTATGCC 5114
QY 1997 AGCTACTGGAGGCTGAGGCAAGAGATCGCTTGAACCCAGAGGTGGAGTTGACATA 2056
Db 5115 AGCTACTGGAGGCTGAGGCAAGAGATCGCTTGAACCCAGAGGTGGAGTTGACATA 5174
QY 2057 AGCTGAGATGTCGCGCTGCACTCCAGCTGGCGCAAGAGCAAGACTCTGTCTAGAA 2116
Db 5175 AGCTGAGATGTCGCGCTGCACTCCAGCTGGCGCAAGAGCAAGACTCTGTCTAGAA 5234
QY 2117 AAAAAAAG 2176
Db 5235 AAAAAAAG 5294
QY 2177 AAGCATTCAG 2236
Db 5295 AAGCATTCAG 5354
QY 2237 TAATTTGCTCTTTGGTGTCTATTTGCTTAACTGCTTGAAGTGAAGAGAGAG 2296
Db 5355 TAATTTGCTCTTTGGTGTCTATTTGCTTAACTGCTTGAAGTGAAGAGAGAG 5414
QY 2297 TTGAGATTAACCATTCCTGTCGCTGTTATTAATTAATTAATTAATTAATTAAT 2356
Db 5415 TTGAGATTAACCATTCCTGTCGCTGTTATTAATTAATTAATTAATTAATTAAT 5474
QY 2357 AGATTCAG 2416
Db 5475 AGATTCAG 5534
QY 2417 TGTAAAGAGTAAATTTGCTTCAATGATTTCAATTAATTAATTAATTAATTAAT 2476
Db 5535 TGTAAAGAGTAAATTTGCTTCAATGATTTCAATTAATTAATTAATTAATTAAT 5594
QY 2477 CTATGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2536
Db 5595 CTATGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5654
QY 2537 TTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2596
Db 5655 TTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5714
QY 2597 AGTAAATGTCCTAG 2656
Db 5715 AGTAAATGTCCTAG 5774
QY 2657 TGTCTTGTGCTCACTTCAATCAATTAATTAATTAATTAATTAATTAATTAAT 2716
Db 5775 TGTCTTGTGCTCACTTCAATCAATTAATTAATTAATTAATTAATTAATTAAT 5834
QY 2717 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2776
Db 5835 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5894
QY 2777 TATGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2797
Db 5895 TATGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5915

RESULT 4
BD016860
LOCUS BD016860 11204 bp DNA linear part 27-AUG-2002
DEFINITION Novel cytidine deaminase.

ACCESSION BD016860
VERSION BD016860.1 GI:22558036
KEYWORDS JP 2001245669-A/33.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
Horiuchi, T. and Muramatsu, M.
Novel cytidine deaminase
Patent: JP 2001245669-A 33 11-SEP-2001;
JOURNAL JAPAN TOBACCO INC./TASUKU HONDO
OS Homo sapiens (human)
PN JP 2001245669-A/33
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092991
PI TASUKU HONDO/MASAMICHI MURAMATSU
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P1/06, A61P13/12,
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12N5/78, C12P21/02, C12P21/08, C12N1/21, C12R1/19, PC
(C12N5/10, C12R1/91), C12N15/00, C12N5/00, (C12N5/00, C12R1/91) CC
FH Key Location/Qualifiers
FEATURES
source 1..11204
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 77.2%; Score 2174.6; DB 6; Length 11204;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 617 CTTTGGCCCTGATGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 676
Db 8950 CTACAGCCCTGATGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 9009
QY 677 TAGCACTTCAGAGAGAGTCAACACAGATGAATATCTGCTGAAGACAGAGATGATAA 736
Db 9010 TAGCACTTCAGAGAGAGTCAACACAGATGAATATCTGCTGAAGACAGAGATGATAA 9069
QY 737 AACGCTCTCAAGTCTCTCTGTTTATCTTCACTCACTCACTTCTTAGAGTTACA 796
Db 9070 AACGCTCTCAAGTCTCTCTGTTTATCTTCACTCACTCACTTCTTAGAGTTACA 9129
QY 797 GAAAAATATTTATATATGAGTCTTTTAAAGATCTATGCTTGAAGATGAGAGAAC 856
Db 9130 GAAAAATATTTATATATGAGTCTTTTAAAGATCTATGCTTGAAGATGAGAGAAC 9189
QY 857 ACAGGCTGGCGAGGAGCGTGGCAATGAGTGAATGAGTGAATGAGTGAATGAGTGAAT 916
Db 9190 ACAGGCTGGCGAGGAGCGTGGCAATGAGTGAATGAGTGAATGAGTGAATGAGTGAAT 9249
QY 917 CTGGGAATTAACAGACTGACAGACTGCGGAGCATCTTAAAGTCTCAAGTTTCTATGA 976
Db 9250 CTGGGAATTAACAGACTGACAGACTGCGGAGCATCTTAAAGTCTCAAGTTTCTATGA 9309
QY 977 CTTTACGTGATGAGAGAGAGAGTGAATCTTAAAGATGAGTGAAGATGAGATCAATG 1036
Db 9310 CTTTACGTGATGAGAGAGAGAGTGAATCTTAAAGATGAGTGAAGATGAGATCAATG 9369
QY 1037 TTTTATATCAACATCTTATATATGATGATGATGATGATGATGATGATGATGATGAT 1096
Db 9370 TTTTATATCAACATCTTATATATGATGATGATGATGATGATGATGATGATGATGAT 9429
QY 1097 GATTTTCTATCTTTTCCCTTGAAGTTACTTCAAGTATCAACAACTCTTCAACAG 1156
Db 9430 GATTTTCTATCTTTTCCCTTGAAGTTACTTCAAGTATCAACAACTCTTCAACAG 9489
QY 1157 CCATGATCTATGAGAGCTCTTAAATGAGAGTATCTGGGTGATGATGATGATGATGAT 1216

D 9490 CCAATGATCTATAGACCTCTCTAATGAGATCTGGTATTTGTGACCCCAACCACTCTC 9549

Q 1217 TCCAAAGCATTTAATATCAATCATGCGCTGTATGTTTATTCAGACAGAAAGATGTTTAA 1276

D 9550 TCCAAAGCATTTAATATCAATCATGCGCTGTATGTTTATTCAGACAGAAAGATGTTTAA 9609

Q 1277 TGTATTGACAAAAGAAAGATTGTTATGGGTGGGATGAGGTATAGACATGATGGTCA 1336

D 9610 TGTATTGACAAAAGAAAGATTGTTATGGGTGGGATGAGGTATAGACATGATGGTCA 9669

Q 1337 CTTCAGACTCTTAAATAGAACTTAAATAGGAGAGAGAGCTGTGAACAGACACCC 1396

D 9670 CTTCAGACTCTTAAATAGAAAGATTAAATAGGAGAGAGCTGTGAACAGACACCC 9729

Q 1397 TAAATATGGGTGATGCTGATAGTAAATCTTCTGAAACGCAAACTCTTTTAAAGAA 1456

D 9730 TAAATATGGGTGATGCTGATAGTAAATCTTCTGAAACGCAAACTCTTTTAAAGAA 9789

Q 1457 GTCCCTAATTTAGAAACACCCACAACTTCAATATCATATTTAGCAAACTTTGAGAG 1516

D 9790 GTCCCTAATTTAGAAACACCCACAACTTCAATATCATATTTAGCAAACTTTGAGAG 9849

Q 1517 AAGTGTCTGAATGTTGGGAGAGAAATCTATTTGGCTCTGTTGGTCTTTCATCTCA 1576

D 9850 AAGTGTCTGAATGTTGGGAGAGAAATCTATTTGGCTCTGTTGGTCTTTCATCTCA 9909

Q 1577 GAAATGCAATCAGTCAAGCTTGTCTACATTTTGTATGTGTGTATGCTTCTCCCAAG 1636

D 9910 GAAATGCAATCAGTCAAGCTTGTCTACATTTTGTATGTGTGTATGCTTCTCCCAAG 9969

Q 1637 GTATATTAATCTATATAGAGAGTGTGACAAAACAGATGATTAAGCTGCAACCGTGGC 1696

D 9970 GTATATTAATCTATATAGAGAGTGTGACAAAACAGATGATTAAGCTGCAACCGTGGC 10029

Q 1697 ACAAGCTCATATGTTCTAGTCTGTGGAGGTTGAGAGAGAGATGGCTTGAACAAGT 1756

D 10030 ACAAGCTCATATGTTCTAGTCTGTGGAGGTTGAGAGAGAGATGGCTTGAACAAGT 10089

Q 1757 GTTCAAGGCGAGCTTGGGCAACATACAGATCTCTCTCAAAAAAAAAAAAAAAAAA 1816

D 10090 GTTCAAGGCGAGCTTGGGCAACATACAGATCTCTCTCTCAAAAAAAAAAAAAAAAAA 10149

Q 1817 AAGAAAGAGAGAGGCGCGGCGGTGTGCTCAAGCTGTAACTTCCAGCACTTTGGAGGC 1876

D 10150 AAGAAAGAGAGAGGCGCGGCGGTGTGCTCAAGCTGTAACTTCCAGCACTTTGGAGGC 10209

Q 1877 CGAGCCGAGCGATCACTGTGTGAGAGTGTGACACAGCTGCGCAACATGGCAAAA 1936

D 10210 CGAGCCGAGCGATCACTGTGTGAGAGTGTGACACAGCTGCGCAACATGGCAAAA 10269

Q 1937 CCCCCTCTCTACTCAAAATGCAAAATTAAGCAAGCGGTGTAGCAAGCACTGTATCCC 1996

D 10270 CCCCCTCTCTACTCAAAATGCAAAATTAAGCAAGCGGTGTAGCAAGCACTGTATCCC 10329

Q 1997 AGTATATTGGAGGCTGAGAGAGAGAAATCGCTTGAACCCAGAGAGTGGAGTTGCAATA 2056

D 10330 AGTATATTGGAGGCTGAGAGAGAGAAATCGCTTGAACCCAGAGAGTGGAGTTGCAATA 10389

Q 2057 AGCTGAGATCGTGCCTTGCATCCAGCTGCGGCGCAAGAGACAGACCTGTCTCAAAA 2116

D 10390 AGCTGAGATCGTGCCTTGCATCCAGCTGCGGCGCAAGAGACAGACCTGTCTCAAAA 10449

Q 2117 AAAAAAAAAAAAAAAAAAGAGAGAGAGAGAAAGAAATATTGGAGAGAGAGATGGGG 2176

D 10450 AAAAAAAAAAAAAAAAAAGAGAGAGAGAGAAAGAAATATTGGAGAGAGAGATGGGG 10509

Q 2177 AAGCATTTGAGAGAAATTTGTCTTTATCCAAATAATTTAGAGCCCAATAGAGATCCC 2236

D 10510 AAGCATTTGAGAGAAATTTGTCTTTATCCAAATAATTTAGAGCCCAATAGAGATCCC 10569

Q 2237 TATTGTCTCTTTGGTGTCTATTGTTGCTTAAACAGTCTTTTGAACGTAGAGAAAAATA 2296

D 10570 TATTGTCTCTTTGGTGTCTATTGTTGCTTAAACAGTCTTTTGAACGTAGAGAAAAATA 10629

Q 2297 TTCAGAAATPACCATATCCCTGTGCGGTATTAATCTAGCAACCCCTGGAAATGAGATGAGC 2356

D 10630 TTCAGAAATPACCATATCCCTGTGCGGTATTAATCTAGCAACCCCTGGAAATGAGATGAGC 10689

Q 2357 AGATCCACAGGAAACTTGAATGCAACAGTCTTATTTTAACTTATTTAGACATAAGTT 2416

D 10690 AGATCCACAGGAAACTTGAATGCAACAGTCTTATTTTAACTTATTTAGACATAAGTT 10749

Q 2417 TGTAAAGAGTTAAATTTGTTACTTCAATGATATTCATTATTAATTTATTTTGGCT 2476

D 10750 TGTAAAGAGTTAAATTTGTTACTTCAATGATATTCATTATTAATTTATTTTGGCT 10809

Q 2477 CTATGATTTTATTTAATGATATTCCTTCTTCTGATATTTGAAATGGAGCTCAAGC 2536

D 10810 CTATGATTTTATTTAATGATATTCCTTCTTCTGATATTTGAAATGGAGCTCAAGC 10869

Q 2537 TTCATTAATTTATTAATCTTGAATGATTTCTAATPACAGGTATGATGTAACATTGC 2596

D 10870 TTCATTAATTTATTAATCTTGAATGATTTCTAATPACAGGTATGATGTAACATTGC 10929

Q 2597 AGAATGGTGTACAGACCATTTCTCTGATTTTAACTTTATGACAGCAATT 2656

D 10930 AGAATGGTGTGTACAGACCATTTCTCTGATTTTAACTTTATGACAGCAATT 10989

Q 2657 TGTCTTGGCTCACTTTCATCAGTTAAATPAAATPAAATTTTGGAGCTGTGAAG 2716

D 10990 TGTCTTGGCTCACTTTCATCAGTTAAATPAAATPAAATTTTGGAGCTGTGAAG 11049

Q 2717 ATAAATATACCAATPAAATTAATTAAGATTTTATGACATTAATTAATTAATTAATTCG 2776

D 11050 ATAAATATACCAATPAAATTAATTAAGATTTTATGACATTAATTAATTAATTAATTCG 11109

Q 2777 TATGATGAATPAACTTGAA 2797

D 11110 TATGATGAATPAACTTGAA 11130

RESULT 5
AB040430 11204 bp DNA linear pri 03-OCT-2000
LOCUS
DEFINITION Homo sapiens AID gene for activation-induced cytidine deaminase,
complete cds.
ACCESSION
AB040430
VERSION
AB040430.1 GI:9988407
KEYWORDS
AID; activation-induced cytidine deaminase.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (sites)
Muto, T., Muramatsu, M., Tanikawa, M., Kinoshita, K. and Honjo, T.
Isolation, tissue distribution, and chromosomal localization of the
human activation-induced cytidine deaminase (AID) gene
Genomics 68 (1), 85-88 (2000)
2 (sites)
Muto, T., Muramatsu, M., Tanikawa, M., Kinoshita, K. and Honjo, T.
Activation-induced cytidine deaminase (AID) deficiency causes the
autosomal recessive form of the Hyper-IgM syndrome (HIGM2)
Cell 1102 (5), 565-575 (2000)
TITLE
JOURNAL
MEDLINE
PUBMED
11007475
3 (bases 1 to 11204)
Muto, T., Muramatsu, M., Tanikawa, M., Kinoshita, K. and Honjo, T.
Direct Submision
Submitted (18-MAR-2000) Tasuku Honjo, Kyoto University, Department
of Medical Chemistry, Faculty of Medicine, Yoshida, Sakyo-ku,

FEATURES
Kyoto, Kyoto 606-8501, Japan (E-mail: honjo@mfour.med.kyoto-u.ac.jp)
Tel: 81-75-753-4371 (ex. 4371), Fax: 81-75-753-4388
Location/Qualifiers

source

1. 11204

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
```

```

/obj_xref="xref:9606"
join(521. .528,6280. .6427,7807. .8077,8371. .8486,8956. .9009)

```

Db 10750 TGTAAAGAGTAAATTTCTACTGATGATTCATTTATATTTATTTATTTGGCT 10809

QY 2477 CTATGATTTTATTAACGATTTCTTTCTGATATATTTGAATGAGTCTCAAGC 2536

Db 10810 CTATGATTTTATTAACGATTTCTTTCTGATATATTTGAATGAGTCTCAAGC 10869

QY 2537 TTTCATTAATTTATTAACGATTTCTTTCTGATATATTTGAATGAGTCTCAAGC 2596

Db 10870 TTTCATTAATTTATTAACGATTTCTTTCTGATATATTTGAATGAGTCTCAAGC 10929

QY 2597 AGTAATGCTGCTACGAGACCATTTCTGATTTTATTAACGATTTCTCAAGCAAT 2656

Db 10930 AGTAATGCTGCTACGAGACCATTTCTGATTTTATTAACGATTTCTCAAGCAAT 10989

QY 2657 TGGTTCTGCTCATCTTCAATGATTTATTAATGATTTATTAATTTTGAAGCTGTGAAG 2716

Db 10990 TGGTTCTGCTCATCTTCAATGATTTATTAATGATTTATTAATTTTGAAGCTGTGAAG 11049

QY 2717 ATAAATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2776

Db 11050 ATAAATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11109

QY 2777 TATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2797

Db 11110 TATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11130

RESULT 6
AC092184
LOCUS
DEFINITION Homo sapiens 12 BAC RP11-438L7 (Roswell Park Cancer Institute Human BAC library) complete sequence.
AC092184 AC013443
VERSION
KEYWORDS
SOURCE HTG.
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Zukaroyta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 71132)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alshrocks,S.L., Amaratunga,H.C., Are,J.R., Ayele,M., Banks,T., Barbieri,U., Benton,J., Bimonte,K., Blankenburg,K., Bonini,D., Bouck,U., Bowie,S., Brileva,M., Brown,M., Brown,M., Bryant,N.P., Buhay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Geo,U., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,K., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B., Homst,F., Howard,S., Huber,U., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Kervah,U., Kovar,C., Kratovic,U., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,U., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Loulesged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Maronde,I., Martin,R., Martindale,A., Martinez,E., Massey,B., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okumotu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Repen,Y.,

Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shoohtari,N., Sisson,I., Sodergren,E., Sonakke,I., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umanu,K., Vasquez,L., Vera,V., Villalón,D., Vismán,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Zuchelapatti,R., Weinstein,G., and Gibbs,R.

TITLE
JOURNAL
Unpublished
2 (bases 1 to 71132)
Worley,K.C.
REFERENCE
AUTHORS
Direct Submission
Submitted (25-JUN-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 71132)
Worley,K.C.
REFERENCE
AUTHORS
Direct Submission
Submitted (18-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 71132)
Worley,K.C.
REFERENCE
AUTHORS
Direct Submission
Submitted (25-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 71132)
Worley,K.C.
REFERENCE
AUTHORS
Direct Submission
Submitted (12-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 25, 2002 this sequence version replaced gi:20901754.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

COMMENT
JOURNAL
TITLE
AUTHORS
JOURNAL

CHROM LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu/8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

```
FEATURES
Source      Location/Qualifiers
1. 71132
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-438L7"
complement(1..1448)
/note="overlaps bases 1..1448 of clone AC092490"
/function="clone overlap"
187..408
/standard_name="57233"
439..560
/standard_name="92005"
774..881
/rpt_family="MIR"
903..1190
/rpc_family="AluSg"
1191..1213
/rpc_family="AT_rich"
1591..1807
/standard_name="6612"
1744..1819
/standard_name="8198"
1966..2264
/rpc_family="AluSg"
3296..3328
/rpc_family="(TTTC)n"
4087..4249
/rpc_family="AluSg"
4360..4654
/rpc_family="AluY"
complement(5140..5262)
/rpc_family="FLAM_C"
5275..5304
/rpc_family="AT_rich"
5305..5563
/rpc_family="AluDo"
5564..5599
/rpc_family="AT_rich"
5644..5733
/rpc_family="MSTD"
5734..6012
/rpc_family="AluSg"
6020..6336
/rpc_family="AluSx"
6337..6629
/rpc_family="MSTD"
6630..6650
/rpc_family="(CAA)n"
complement(6711..6817)
/rpc_family="MIR"
complement(7033..7327)
/rpc_family="AluSx"
7607..7745
/rpc_family="FLAM_A"
7980..8047
/rpc_family="MIR"
complement(8815..9101)
/rpc_family="AluSg"
complement(11227..12175)
/rpc_family="LTR5"
12625..13900
/standard_name="87432"
complement(13547..13664)
/rpc_family="LIMB5"
complement(13665..13979)
```

```
repeat_region      /rpt_family="AluY"
complement(13980..14193)
repeat_region      /rpt_family="LIMB5"
14622..14924
repeat_region      /rpt_family="AluY"
complement(15813..17136)
repeat_region      /rpt_family="LIMB8"
complement(17272..17395)
repeat_region      /rpt_family="FLAM_C"
17426..17575
repeat_region      /rpt_family="MIR"

Query Match      77.2%; Score 2174.6; DB 9; Length 71132;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

617 CTTTGGCCCTGATGAGTTGATGACTTACGAGCAGCATTTGCTACTTTGGACTTGA 676
|||
43493 CTACAGCCCTGATGAGTTGATGACTTACGAGCAGCATTTGCTACTTTGGACTTGA 43552
|||
677 TAGCACTTCAGGAATGTCACACAGATGAATATCTCTGCTGAAGACAGTGATPAA 736
|||
43553 TAGCACTTCAGGAATGTCACACAGATGAATATCTCTGCTGAAGACAGTGATPAA 43612
|||
737 AACAGTCTTCAAGTCTCTCTGTTTATTTCTTCAACTCTCACTTTCTTAGGTTTCA 796
|||
43613 AACAGTCTTCAAGTCTCTCTGTTTATTTCTTCAACTCTCACTTTCTTAGGTTTCA 43672
|||
797 GAAAAATTTTATATGAGCTCTTTAAAGATCTATGCTTGAATPAGAGAGAAC 856
|||
43673 GAAAAATTTTATATGAGCTCTTTAAAGATCTATGCTTGAATPAGAGAGAAC 43732
|||
857 ACAGTCTGCGACGAGACGTGCTGCAATTGTCAGTTTGAATGACATTGCTCCCTA 916
|||
43733 ACAGTCTGCGACGAGACGTGCTGCAATTGTCAGTTTGAATGACATTGCTCCCTA 43792
|||
917 CTGGGAATPACAGACTGTCAGACCTGGAGCATCTTAAAGTGTCAAGTTTCTATGA 976
|||
43793 CTGGGAATPACAGACTGTCAGACCTGGAGCATCTTAAAGTGTCAAGTTTCTATGA 43852
|||
977 CTTTATGATGAGAGAGAGAGAGAGATCTTAAAGATGAGAGAGATGCAAAAG 1036
|||
43853 CTTTATGATGAGAGAGAGAGAGATCTTAAAGATGAGAGAGATGCAAAAG 43912
|||
1037 TTTTATATCAACATCTTATTTATTTGATTGATTGATTGATTGATTGATTGATTGATT 1096
|||
43913 TTTTATATCAACATCTTATTTATTTGATTGATTGATTGATTGATTGATTGATTGATT 43972
|||
1097 GATTTTCTATTTCTTTTCCCTTACGTTTCACTTCAATGACAAACTCTTCATCAGG 1156
|||
43973 GATTTTCTATTTCTTTTCCCTTACGTTTCACTTCAATGACAAACTCTTCATCAGG 44032
|||
1157 CCATGATCTATGAGACCTCTATGAGAGATCTGAGTATTTGATGAGAGAGAGAGAGAG 1216
|||
44033 CCATGATCTATGAGACCTCTATGAGAGATCTGAGTATTTGATGAGAGAGAGAGAGAG 44092
|||
1217 TCGAAGATTAATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1276
|||
44093 TCGAAGATTAATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 44152
|||
1277 TGTTTGACAAAGAGATTGATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1336
|||
44153 TGTTTGACAAAGAGATTGATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 44212
|||
1337 CTTCAAGTACTTTAATTAAGAGATCTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1396
|||
44213 CTTCAAGTACTTTAATTAAGAGATCTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 44272
|||
1397 TAAATAGGAGTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1456
|||
44273 TAAATAGGAGTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 44332
|||
1457 GTCCCTAATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1516
|||
```

```

Db      44333 GTCCCAATTTTAAACACCCCAAACTTCAATATCAATATAGCAACAATTGGAAG 44392
Qy      1517 AAGTCTTGATGATGGGAGAGAAAATCTATGGCTGTGGGTCTCTCAATCCA 1576
Db      44393 AAGTCTTGATGATGGGAGAGAAAATCTATGGCTGTGGGTCTCTCAATCCA 44452
Qy      1577 GAATGCCAATCAGGTCAAGGTTGCTAATTTGTATGTGTGTGCTTCCCAAG 1636
Db      44453 GAATGCCAATCAGGTCAAGGTTGCTAATTTGTATGTGTGTGCTTCCCAAG 44512
Qy      1637 GTATATTAATATATAGAGAGTTGTGACAAAACAGATGATPAAGCTGCCAG 1696
Db      44513 GTATATTAATATATAGAGAGTTGTGACAAAACAGATGATPAAGCTGCCAG 44572
Qy      1697 ACAGGCTCATAGTCTTAGGCTTGGAGTTGAGAGAGAGATGGCTTGAACAGGT 1756
Db      44573 ACAGGCTCATAGTCTTAGGCTTGGAGTTGAGAGAGAGATGGCTTGAACAGGT 44632
Qy      1757 GTTCAAGCCAGCCTGGGCAACATAACAGATCTGTCTCAAAAAA 1816
Db      44633 GTTCAAGCCAGCCTGGGCAACATAACAGATCTGTCTCAAAAAA 44692
Qy      1817 AAGAAAGAGAGAGGCGCGGCGTGTGCTCAAGCTGTATCCAGCATTTGGAGGC 1876
Db      44693 AAGAAAGAGAGAGGCGCGGCGTGTGCTCAAGCTGTATCCAGCATTTGGAGGC 44752
Qy      1877 CGAGCCGGCGGATCACTGTGTGCTCAGAGTTTGAGACCAAGCTGGCAAA 1936
Db      44753 CGAGCCGGCGGATCACTGTGTGCTCAGAGTTTGAGACCAAGCTGGCAAA 44812
Qy      1937 CCCGCTGTACTCAAAATGCCAAAATTAGCCAGGCGGTGTGAGGAGCCTGTATCCC 1996
Db      44813 CCCGCTGTACTCAAAATGCCAAAATTAGCCAGGCGGTGTGAGGAGCCTGTATCCC 44872
Qy      1997 AGTACTTGGAGGCTGAGGAGAGAGAAATCGCTTGAACCCAGAGAGTGAAGTTG 2056
Db      44873 AGTACTTGGAGGCTGAGGAGAGAGAAATCGCTTGAACCCAGAGAGTGAAGTTG 44932
Qy      2057 AGTGAATGTCGCGCTTGCCTCACTCCAGCTGGGCGACAAGACCAAGCTGTCTC 2116
Db      44933 AGTGAATGTCGCGCTTGCCTCACTCCAGCTGGGCGACAAGACCAAGCTGTCTC 44992
Qy      2117 AAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2176
Db      44993 AAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 45052
Qy      2177 AAGCATTTGAGAGAAATGTGCTTATCCAAACAAATGTAAGAGCCAAATAGGATCCC 2236
Db      45053 AAGCATTTGAGAGAAATGTGCTTATCCAAACAAATGTAAGAGCCAAATAGGATCCC 45112
Qy      2237 TATTTGCTCTTTGGTGTCTATTTGCTTAACTCAACTGTCTTGAACAGTGA 2296
Db      45113 TATTTGCTCTTTGGTGTCTATTTGCTTAACTCAACTGTCTTGAACAGTGA 45172
Qy      2297 TTCAGAAATACATATCCCTGTGCGCTTATTAAGTCAAGCCCTTGAATGAATGAGC 2356
Db      45173 TTCAGAAATACATATCCCTGTGCGCTTATTAAGTCAAGCCCTTGAATGAATGAGC 45232
Qy      2357 AGATCCACAGAGAACTTGAATGCAACTGTCTTATTTAATCTTATTTGAACAT 2416
Db      45233 AGATCCACAGAGAACTTGAATGCAACTGTCTTATTTAATCTTATTTGAACAT 45292
Qy      2417 TGTAAAGAGTAAATATTTACTTCAATGATCAATTAATTTATTTATTTGGGT 2476
Db      45293 TGTAAAGAGTAAATATTTACTTCAATGATCAATTAATTTATTTATTTGGGT 45352
Qy      2477 CTAAATGATTTTATTAACATGATTTCTTTTGTGATATTAATGAATGAGTCAAGC 2536
Db      45353 CTAAATGATTTTATTAACATGATTTCTTTTGTGATATTAATGAATGAGTCAAGC 45412
Qy      2537 TTCATTAATTTAATCTTTAGAAATGATTTCTAATPAACAGATGATTTGAACAT 2596

```

```

Db      45413 TTCATTAATTTAATCTTTAGAAATGATTTCTAATPAACAGATGATTTGAACAT 45472
Qy      2597 AGTAAATGATGCTAGAGAGCCATTTCTCTGATTTTATAGTAACTTTATGACGCAAT 2656
Db      45473 AGTAAATGATGCTAGAGAGCCATTTCTCTGATTTTATAGTAACTTTATGACGCAAT 45532
Qy      2657 TGCTTGGCTCACTTCAATGATTTAATTAATTAATTAATTTTGAAGCTGGAAG 2716
Db      45533 TGCTTGGCTCACTTCAATGATTTAATTAATTAATTAATTTTGAAGCTGGAAG 45592
Qy      2717 ATAAATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2776
Db      45593 ATAAATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 45652
Qy      2777 TATGATGAATTAATCTGAAA 2797
Db      45653 TATGATGAATTAATCTGAGA 45673

RESULT 7
BD016840
LOCUS      BD016840          2172 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Novel cytidine deaminase.
ACCESSION BD016840
VERSION    BD016840.1 GI:22558016
KEYWORDS   UP 2001245669-A/13.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2172)
AUTHORS   Honjo, T. and Muramatsu, M.
TITLE      Novel cytidine deaminase
JOURNAL    Patent: JP 2001245669-A 13 11-SEP-2001;
JOURNAL    JAPAN TOBACCO INC. TASUKU HONJO
COMMENT    OS Homo sapiens (human)
           PN JP 2001245669-A/13
           PD 11-SEP-2001
           PF 28-MAR-2000 JP 2000092981
           PI TASUKU HONJO, MASAMICHI MURAMATSU
           PC C12N15/00, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
           PC A61P17/00,
           PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
           C12N1/21,
           PC C12N5/10, C12N9/78, C12P21/02, C12P21/08// (C12N1/21, C12R1/19), PC
           (C12N5/10, C12R1/91), C12N15/00, C12N5/00, (C12N5/00, C12R1/91) CC
           FH Key
FEATURES
SOURCE
location/Qualifiers
1..2172
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 77.1%; Score 2172; DB 6; Length 2172;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	86	CTGGCCAGGGAGCTGCTCCTCAATGGGTGAGATGTTGAAATGACAATGTCCTCCATCGGGAG	922
Db	241	CTGGCCAGGGAGCTGCTCCTCAATGGGTGAGATGTTGAAATGACAATGTCCTCCATCGGGAG	300
QY	923	ATPACAGAACTGCAGAGACTGGAGACATCTTAAAGTGTCAACGTTTCTATGACTTTTAA	982
Db	301	ATPACAGAACTGCAGAGACTGGAGACATCTTAAAGTGTCAACGTTTCTATGACTTTTAA	360
QY	983	GGTAGATGAGACGAAAGGTGATCCTTAAAGCATGTGTGAGAGATGAAATGTTTAA	1042
Db	361	GGTAGATGAGACGAAAGGTGATCCTTAAAGCATGTGTGAGAGATGAAATGTTTAA	420
QY	1043	TATCAACATCCTTTATTTATTTGATTCAATTGAGTTACAATGGTGTATGATGATATTT	1102
Db	421	TATCAACATCCTTTATTTATTTGATTCAATTGAGTTACAATGGTGTATGATGATATTT	480
QY	1103	TCTATTCCTTCCCTTGACGTTTACTTCAAGTAAACAACATCTTCCATCGAGCATGA	1162
Db	481	TCTATTCCTTCCCTTGACGTTTACTTCAAGTAAACAACATCTTCCATCGAGCATGA	540
QY	1163	TCTATGAGACTCCTTATGAGAGTATCTGGGTGATTGTGACCCCAACCATCTTCCAA	1222
Db	541	TCTATGAGACTCCTTATGAGAGTATCTGGGTGATTGTGACCCCAACCATCTTCCAA	600
QY	1223	GCAATTATATCCAAATCAGCGCTGTATGTTTAACTAGCGAAGCATGTTTATATGTTTG	1282
Db	601	GCAATTATATCCAAATCAGCGCTGTATGTTTAACTAGCGAAGCATGTTTATATGTTTG	660
QY	1283	TACAAAGAAATGTTTATGGTGGGGATGGAGGTATAGACATGATGTCACCTTGA	1342
Db	661	TACAAAGAAATGTTTATGGTGGGGATGGAGGTATAGACATGATGTCACCTTGA	720
QY	1343	GCTACTTATATTAAGATCTTAAATGGGCGAGAGACTGTGAACAAGACACCTTATPA	1402
Db	721	GCTACTTATATTAAGATCTTAAATGGGCGAGAGACTGTGAACAAGACACCTTATPA	780
QY	1403	TGGTGTGATGTCTGAGTAGCAAACTTCTTGAAACGCAACTTTTAAAGAACTCCT	1462
Db	781	TGGTGTGATGTCTGAGTAGCAAACTTCTTGAAACGCAACTTTTAAAGAACTCCT	840
QY	1463	AATTTGAAACCCACAACTTCAACATATATTAATTAGCAACATTTGAAAGAAATGG	1522
Db	841	AATTTGAAACCCACAACTTCAACATATATTAATTAGCAACATTTGAAAGAAATGG	900
QY	1523	CTTGAATGTTGGGAGAGGAAATCTATGGCTCGTGGGTCTCTTCACTCAGAAATG	1582
Db	901	CTTGAATGTTGGGAGAGGAAATCTATGGCTCGTGGGTCTCTTCACTCAGAAATG	960
QY	1583	CCAATCAGTCAAGGTTTGCTACATTTGTGTATGTGTGATGCTTCCCAAGGATAT	1642
Db	961	CCAATCAGTCAAGGTTTGCTACATTTGTGTATGTGTGATGCTTCCCAAGGATAT	1020
QY	1643	TAACTTATTAAGAGTTGTGACAAACAGATATTAAGCTCGCAACCTGGCACAGC	1702
Db	1021	TAACTTATTAAGAGTTGTGACAAACAGATATTAAGCTCGCAACCTGGCACAGC	1080
QY	1703	TCAATGTTTACCTCTGGAGAGTTGAGAGGGAGATGGCTTGAACACAGTGTCTAA	1762
Db	1081	TCAATGTTTACCTCTGGAGAGTTGAGAGGGAGATGGCTTGAACACAGTGTCTAA	1140
QY	1763	GGCAGACCTGGGCAACATTAACAAGTCTGTCTCTMAAAAAAAAAAAAAAAAAAGAA	1822
Db	1141	GGCAGACCTGGGCAACATTAACAAGTCTGTCTCTMAAAAAAAAAAAAAAAAAAGAA	1200
QY	1823	GAGAGAGGAGCCGGGCGTGTGCTCAAGCTGTATCCAGCACTTGGAGAGCGAGCC	1882
Db	1201	GAGAGAGGAGCCGGGCGTGTGCTCAAGCTGTATCCAGCACTTGGAGAGCGAGCC	1260
QY	1883	GGGCGGATACCTGTGTGTCAAGAGTTTGAACACAGCTGGCCAAACATGGCAAAACCCGT	1942
Db	1261	GGGCGGATACCTGTGTGTCAAGAGTTTGAACACAGCTGGCCAAACATGGCAAAACCCGT	1320

Oy	1943	TTGTAACCTCAAAATGCAAAAATTTAGCCGAGGGGTGGTAGCAGGACCTGTAAATCCAGACTAC	2002
Db	1321	CTGTAACCTCAAAATGCAAAAATTTAGCCGAGGGGTGGTAGCAGGACCTGTAAATCCAGACTAC	1380
Oy	2003	TTGGGAGGCTGAGCAGAGAGATTCGCTTGAACCCAGAGAGGTGAGGATTGGCATGAAGCTGA	2062
Db	1381	TTGGGAGGCTGAGCAGAGAGATTCGCTTGAACCCAGAGAGGTGAGGATTGGCATGAAGCTGA	1440
Oy	2063	GATGTGCGCGTTGTCATCTCCAGCTGGGGCGACAGAGCAGATCTGTCTCGAAAAAAA	2122
Db	1441	GATGTGCGCGTTGTCATCTCCAGCTGGGGCGACAGAGCAGATCTGTCTCGAAAAAAA	1500
Oy	2123	AAAAAAAAGAGAGAGAGAGAAAAGAAACAATTTTGGAGAGAGAAAGATGGGGAGACAT	2182
Db	1501	AAAAAAAAGAGAGAGAGAGAAAAGAAACAATTTTGGAGAGAGAAAGATGGGGAGACAT	1560
Oy	2183	TGCAAGAGAAATTTGCTTATTCACAACAAAATGTAGAGAGCCAAATTAAGGATTCCTTATTTG	2242
Db	1561	TGCAAGAGAAATTTGCTTATTCACAACAAAATGTAGAGAGCCAAATTAAGGATTCCTTATTTG	1620
Oy	2243	TCCTTTTGGTGTCATTTGTGCTTAAACAACCTGTCTTGAACGTGAGAAAAATATTGAGA	2302
Db	1621	TCCTTTTGGTGTCATTTGTGCTTAAACAACCTGTCTTGAACGTGAGAAAAATATTGAGA	1680
Oy	2303	ATACCCATATCCCTGTGCGCTTATTAACCTAGCAACCTTGCAATGAAGATGAGAGATCC	2362
Db	1681	ATACCCATATCCCTGTGCGCTTATTAACCTAGCAACCTTGCAATGAAGATGAGAGATCC	1740
Oy	2363	ACAGAGAAACCTGGAATGACAAACCTGTCTTATTTTAACTTATGTAATATAGTTGGTGA	2422
Db	1741	ACAGAGAAACCTGGAATGACAAACCTGTCTTATTTTAACTTATGTAATATAGTTGGTGA	1800
Oy	2423	AGAGTTAAAAATTGTACTTCATGATATCATTTATATATTATTTATTTTGGCGTCAATG	2482
Db	1801	AGAGTTAAAAATTGTACTTCATGATATCATTTATATATTATTTATTTTGGCGTCAATG	1860
Oy	2483	ATTTTATTTATTAACATGATTTCCCTTTCTGTGATATATTTGAATTTGAGTCTCAAGCTTCATA	2542
Db	1861	ATTTTATTTATTAACATGATTTCCCTTTCTGTGATATATTTGAATTTGAGTCTCAAGCTTCATA	1920
Oy	2543	AATTATTAACCTTGAATGATTTCTATTAACAACGATGATGTAATTTGTAAACATTTGAGTAAT	2602
Db	1921	AATTATTAACCTTGAATGATTTCTATTAACAACGATGATGTAATTTGTAAACATTTGAGTAAT	1980
Oy	2603	GGTGCTACGAGACCATTTCTTGATTTTGTAGTAACTTTATGACACGACAAATTTGGCTTC	2662
Db	1981	GGTGCTACGAGACCATTTCTTGATTTTGTAGTAACTTTATGACACGACAAATTTGGCTTC	2040
Oy	2663	TGCGTCACTTTCATCATCGTTAAATGAATGAATTAATTTTGGAAAGCTGTGAGATTA	2722
Db	2041	TGCGTCACTTTCATCATCGTTAAATGAATGAATTAATTTTGGAAAGCTGTGAGATTA	2100
Oy	2723	TACCAATTAATAATTAATAAGATTTATTAACAACCTTAAATTAATAATTCAGTATGAT	2782
Db	2101	TACCAATTAATAATTAATAAGATTTATTAACAACCTTAAATTAATAATTCAGTATGAT	2160
Oy	2783	GGAATTAACCTTG 2794	
Db	2161	GGAATTAACCTTG 2172	
RESULT 8			
BC006296			
LOCUS			
DEFINITION			
Homo sapiens activation-induced cytidine deaminase, mRNA (CDNA			
clone MGC:12911 IMAGE:4054915), complete cds.			
ACCESSION			
BC006296			
VERSION			
BC006296.2 GI:33871601			
KEYWORDS			
MGC.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			

REFERENCE
AUTHORS
1 (bases 1 to 1828)
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schlier G.D.,
Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Schetz T.E., Brownstein M.J., Usdin T.B., Teshigahara S.,
Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Muliyil S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Garabito P.H., Richards S.,
Morley K.C., Hale S., Garcia A.M., Gay L.J., Huijck S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettner M., Madan A., Rodriguez S.,
Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y.,
Boulard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalske U., Small D.E.,
Schnerch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 1828)
Strausberg R.
Direct Submission
Submitted (09-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:13623400.
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Louis Straub
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgti.nih.gov
Achter N., Ayele K., Beckstrom-Sternberg S.M., Benjamin B.,
Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S.,
Dietrich N.L., Granite S., Guan X., Gupta J., Haghghi P.,
Hansen N., Ho S.-L., Karlins E., Kwong P., Lurie P., Legaspi R.,
Maduro Q.L., Mastello C., Maskeri B., Mastrlian S.D., McCloskey J.C.,
McDowell J., Pearson R., Stantipop S., Thomas P.J., Touchman J.W.,
Turegoun C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.,
Young A., Zhang L.-H. and Green E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ULNL at: <http://image.llnl.gov>
Series: IRAL Plate: 17 Row: a Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10190699.
location/Qualifiers
1. 1828
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:12911 IMAGE:4054915"
/cistype="Primary B-Cells from Tonsils"
/clone_id="NIH_MGC_48"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1. 1828
/gene="AICDA"
/note="synonyms: AID, HIGM2, CDA2, ARP2"
/db_xref="Locustid:57379"
/db_xref="MIM:605257"
77. 673

ORIGIN
Query Match 64.5%; Score 1818.6; DB 9; Length 1828;
Beet Local Similarity 99.88; Pred. No. 0;
Matches 1821; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
/codon_start=1
/product="activation-induced cytidine deaminase"
/protein_id="AAH06296.1"
/db_xref="GI:13623401"
/db_xref="Locustid:57379"
/translation="MDSLMNRRELYOPFNWAKRRELYCYVKKRSDSTSPST
DFGLRNKNCHVLELRISMDLDPGCTRYVTFMSPYDCAKRPADLRNP
NLSRLITFALYFEDKRAPEGRRLRAGVAILMTFDFYFCWNTFVENHERFK
AMEGLHENSVLRSQRLRIILPEYVDLDPAFTTLGL"
4 GAACCATCAATTAATGAAGTAGATTTTCTGCGCTGAGACTTGACAGGAGCAAGAGA 63
1 GAACCATCAATTAATGAAGTAGATTTTCTGCGCTGAGACTTGACAGGAGCAAGAGA 60
64 CACTTGAGACAACATAGACAGCCTTGTATGATGACCGAGAGAAATTTCTTACCAATT 123
61 CACTTGAGACAACATAGACAGCCTTGTATGATGACCGAGAGAAATTTCTTACCAATT 120
124 CAAAATGTCGCTGGGCTAAGGCTGAGAGCTGAGACCTTCTGCTAGCTAGTAAAG 183
121 CAAAATGTCGCTGGGCTAAGGCTGAGAGCTGAGACCTTCTGCTAGCTAGTAAAG 180
184 GCGTGAAGAGCTGATCCTTTTCTGCTGAGACTTGTATCTTGCAATAGAAAGGCTG 243
181 GCGTGAAGAGCTGATCCTTTTCTGCTGAGACTTGTATCTTGCAATAGAAAGGCTG 240
244 CCAAGTGAATGCTCTTCTCCGCTCATCTGAGAGCTGAGAGCTGAGAGCTGAGAG 303
241 CCAAGTGAATGCTCTTCTCCGCTCATCTGAGAGCTGAGAGCTGAGAGCTGAGAG 300
304 CTACCGGCTCACTGCTTACCTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 363
301 CTACCGGCTCACTGCTTACCTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 360
364 CGACTTCTGAGAGAGGAAACCCCAACCTGATGAGATTTTCAACCGGCGCTTACTT 423
361 CGACTTCTGAGAGAGGAAACCCCAACCTGATGAGATTTTCAACCGGCGCTTACTT 420
424 CTGAGAGACCGCAAGGCTGAGCCGAGGAGGCTGAGGCTGAGCCGAGGAGGCA 483
421 CTGAGAGACCGCAAGGCTGAGCCGAGGAGGCTGAGGCTGAGCCGAGGAGGCA 480
484 AATGCGCATATGACCTTCAAGATTATTTTCTGCTGAGATATTTTGTAGAAAAACA 543
481 AATGCGCATATGACCTTCAAGATTATTTTCTGCTGAGATATTTTGTAGAAAAACA 540
544 TGAAGAATCTTCAAGCTGAGAGGCTGAGATGAAAATTCAGTCTTCTCCAGACA 603
541 TGAAGAATCTTCAAGCTGAGAGGCTGAGATGAAAATTCAGTCTTCTCCAGACA 600
604 GCTTGCGGAGATCTTTTGGCCCTGTATGAGTATGATTTTCAAGAGCCATTTGGTAC 663
601 GCTTGCGGAGATCTTTTGGCCCTGTATGAGTATGATTTTCAAGAGCCATTTGGTAC 660
664 TTTGGACTTTGATGAGACCTTCCAGAGATGTCACACAGATGAAATATCTGCTGAG 723
661 TTTGGACTTTGATGAGACCTTCCAGAGATGTCACACAGATGAAATATCTGCTGAG 720
724 ACAATGATTAATAAAACAGCTCTTCAAGTCTTCTGTTTATTTTCAACTCTTCA 783
721 ACAATGATTAATAAAACAGCTCTTCAAGTCTTCTGTTTATTTTCAACTCTTCA 780
784 CTTAGATTTTACAGAAAAATATTTATATACAGACTCTTTTAAAAAGATCTATGCTTGA 843
781 CTTAGATTTTACAGAAAAATATTTATATACAGACTCTTTTAAAAAGATCTATGCTTGA 840
844 ATAGAGAGAGACACAGGCTGAGCGAGGAGCTGATGATTTGATGATGATGCA 903

Db 841 ATAGAGAGAGACAGGTCTGGCCAGGACGTGCTGCAATGTCAGATTTTGAATGCA 900
Qy 904 ACATTGCCCCCTACTGGGAATTAACAGAACTGCGACCTGGAGCATCTCTAAAGTCTCA 963
Db 901 ACATTGCCCCCTACTGGGAATTAACAGAACTGCGACCTGGAGCATCTCTAAAGTCTCA 960
Qy 964 CGTTTTCATGACTTTAGTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1023
Db 961 CGTTTTCATGACTTTAGTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Qy 1024 AGAGATCAAAATGTTTATATCAACATCTCTTATTTATTTGATTCATTTAGTTAAAGT 1083
Db 1021 AGAGATCAAAATGTTTATATCAACATCTCTTATTTATTTGATTCATTTAGTTAAAGT 1080
Qy 1084 GGTGTTAGTATGATTTTCTATTTCTTTTCCCTTGAAGTTTCACTTTCAAGTAAACAAA 1143
Db 1081 GGTGTTAGTATGATTTTCTATTTCTTTTCCCTTGAAGTTTCACTTTCAAGTAAACAAA 1140
Qy 1144 CTCTTCATGAGGCGCATGATCTATAGACCTCTCTATGAGATTCCTGGGTGATTTGTGAC 1203
Db 1141 CTCTTCATGAGGCGCATGATCTATAGACCTCTCTATGAGATTCCTGGGTGATTTGTGAC 1200
Qy 1204 CCCAAACCATCTCTCCAAAGCATTAATATCAATCATGCGCTGTATGTTTATATCAGGAG 1263
Db 1201 CCCAAACCATCTCTCCAAAGCATTAATATCAATCATGCGCTGTATGTTTATATCAGGAG 1260
Qy 1264 AAGCATGTTTTATGTTTGTACAAAAGAAAGTTGTTAGGATGGAGATGAGATTAAGC 1323
Db 1261 AAGCATGTTTTATGTTTGTACAAAAGAAAGTTGTTAGGATGGAGATGAGATTAAGC 1320
Qy 1324 CAGCATGTCACCTTCAAGCTCTTTATTAAGATCTTTAAATAGGATGGAGAGAGAGCTGT 1383
Db 1321 CAGCATGTCACCTTCAAGCTCTTTATTAAGATCTTTAAATAGGATGGAGAGAGAGCTGT 1380
Qy 1384 GAACAAGACACCTTAATATGAGTGTGATGTCGAGTACCAATCTCTCGAAGAGCAAA 1443
Db 1381 GAACAAGACACCTTAATATGAGTGTGATGTCGAGTACCAATCTCTCGAAGAGCAAA 1440
Qy 1444 CTCTTTTAAAGAGTCCCTAATTTAGAAACACCAAACTTCAATATCATATTTAGCA 1503
Db 1441 CTCTTTTAAAGAGTCCCTAATTTAGAAACACCAAACTTCAATATCATATTTAGCA 1500
Qy 1504 AACATTTGAGAGAGTGTGATGTTGGGAGAGAGAAATCTATTGGCTCTCGTGGG 1563
Db 1501 AACATTTGAGAGAGTGTGATGTTGGGAGAGAGAAATCTATTGGCTCTCGTGGG 1560
Qy 1564 TCTCTTCATCTCGAATATGCAATTCAGTTCAGGTTTGTCACTTTTGTATGTGTGAT 1623
Db 1561 TCTCTTCATCTCGAATATGCAATTCAGTTCAGGTTTGTCACTTTTGTATGTGTGAT 1620
Qy 1624 GCTTCTCCCAAGAGTATTAATCTATATTAAGAGTGTGACAAAACGAATGATTAAGC 1683
Db 1621 GCTTCTCCCAAGAGTATTAATCTATATTAAGAGTGTGACAAAACGAATGATTAAGC 1680
Qy 1684 TGGGAACCGTGACACGCTCATAGTTCAGCTGCTTGGAGGTTTGAAGAGGAGAGATGG 1743
Db 1681 TGGGAACCGTGACACGCTCATAGTTCAGCTGCTTGGAGGTTTGAAGAGGAGAGATGG 1740
Qy 1744 CTTGAAACACAGGTTTCAAGGCGCAGCTGGCAATTAACAGATCTGTCTCTCAAAA 1803
Db 1741 CTTGAAACACAGGTTTCAAGGCGCAGCTGGCAATTAACAGATCTGTCTCTCAAAA 1800
Qy 1804 AAAAAAAAAAAAAAAAAAGAGAGA 1828
Db 1801 AAAAAAAAAAAAAAAAAAAAAA 1825

RESULT 9
BD016828 2440 bp DNA linear PAT 27-AUG-2002
LOCUS Novel cytidine deaminase.
DEFINITION
ACCESSION BD016828
VERSION BD016828.1 GI:22558004

KEYWORDS JP 2001245669-A/1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Honjo T. and Muramatsu M.
TITLE Novel cytidine deaminase
JOURNAL Patent: JP 2001245669-A 1 11-SEP-2001;
JAPAN TOBACCO INC, TASUKU HONDO
OS Mus musculus (mouse)
PN JP 2001245669-A/1
PD 11-SEP-2001
PF 28-MAR-2000 JP 2000092981
PI TASUKU HONDO, NASAMICHI MURAMATSU
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12,
PC A61P17/00,
PC A61P27/02, A61P37/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12N5/78, C12P21/02, C12P21/08//C12N1/21, C12N1/19, PC
(C12N5/10, C12N1/91), C12N15/00, C12N5/00, C12N5/00, C12N1/91) CC
FH Key Location/Qualifiers
FT CDS (93) (689)
FT 5'UTR (1) (92)
FT 3'UTR (690) (2440).
FEATURES
source 1..2440
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 21.4%; Score 603.4; DB 6; Length 2440;
Best Local Similarity 69.4%; Pred. No. 6,4e-108;
Matches 1006; Conservative 0; Mismatches 351; Indels 92; Gaps 10;
Qy 4 GAACCATCATTAATGAGTGAATTTTCTGSCCTGAGACTTGAAGAGGACAG-AAG 62
Db 16 GAAGACACCTTGCTTGAAGCACTTCTTGSCCTTAAGACTTGAAGAGTCAAGAAAG 75
Qy 63 ACATCTGGAACACCATATGAGACAGCTCTGTATGAACGAGAGAAATTCTTTACCAAT 122
Db 76 TCACGCTGAGACCGGATATGACAGCTTCTGTATGAAGAAAGAAATTCTTTACCAAT 135
Qy 123 TCAAAAATGCGCGTGGGCTTAAGGTCGCGGTGAGACCTACCTGTCCTAGTGAAG 182
Db 136 TCAAAAATGTCGCGTGGGCTTAAGGTCGCGGTGAGACCTACCTGTCCTAGTGAAG 195
Qy 183 GGCCTGACAGTGTACATCTTTTCACTGAGCTTGTATCTTCCCAATTAAGACGGCT 242
Db 196 GGAAGATATGCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 255
Qy 243 GCGACGTGAATGCTCTTCTCGCTACATCTCGGACTGGAACCTTAACCTTGGCGCT 302
Db 256 GCGACGTGAATGCTCTTCTCGCTACATCTCGGACTGGAACCTTGGCGCTGCTGCT 315
Qy 303 GCTAACCGGTCACCTGTTTCACTCTGAGACCGCTGCTACAGACTGTCGCCGCAAGTGG 362
Db 316 GTTACCGGTCACCTGTTTCACTCTGAGACCGCTGCTACAGACTGTCGCCGCAAGTGG 375
Qy 363 CCGACTTTCGAGAGGAAACCCCACTGAGTCTGAGATCTTCAACGCGGCTCTACT 422
Db 376 CTGAGTTTCTGAGATGGAACCTTAACCTGAGTCTGAGATCTTCAACGCGGCTCTACT 435
Qy 423 TCTGTGAGACCGCAAGGCTGAGCCGAGGCGCTGCGCGCTGCAACCGCGCGGCTGC 482
Db 436 TCTGTGAGACCGCAAGGCTGAGCCGAGGCGCTGCGCGCTGCAACCGCGCGGCTGC 495
Qy 483 AATATGCAATCAGACCTTCAAGATTTATTTTACTGCTGGAATATTTTGTGAAAGAAC 542
Db 496 AGATCGGATCAATGACCTTCAAGACTATTTTATGCTGGAATATCATTGTGAAAGAAC 555
Qy 543 ATGAAACATTTCAAGCCTGGAGGCGTGATGAATTCAGTTGCTCTCTCAGAC 602

```

Db      556 GTAAAGAACTTCAAGCTGGGAAGGGCTACATAAAATCTGTCGGGCTACCAAGC 615
Qy      603 AGTTGGGCGGACCTTTTGGCCCTGTATGAGGTATGACTTACGAGCGCATTTTCGTA 662
Db      616 AACTTGGGCGGACCTTTTGGCCCTGTATGAGGTATGACTTACGAGCGCATTTTCGTA 675
Qy      663 CTTCGGGACTTTGATGCACTTCGAGGATGTCACAGATGAAATATCTCTGTGA 722
Db      676 TGTGGGATTTTAAAGCACTCTCTGAAATGTCAAGTGAATTTCT---CTGA 732
Qy      723 GACAGTGAATAAAAAACAGTCTTCAAGTCTCTGTTTATTTCTTCACTCACTT 782
Db      733 GAACTGGAATGAAAAACAACCTTCA--ACTACAGTTTCTTCTTAAAGTACACTT 790
Qy      783 TCTTAAGATTACAGAAAAAATATTAATATACGACTTTTAAAAAGATTTATGCTTGA 842
Db      791 TTATAGTGTAGGGGAAA-----TTATAGACTTTTAAAAAATCTTGAAGCTGCAC 843
Qy      843 AATAGAAAGAACACAGGCTGCGCAGGAGACTGCTCAATTGGTCAGTTTGAATGC 902
Db      844 AGGACCGCCAGACCAATGATGTAAC-----TGAGCTTGCTGTGC 882
Qy      903 AACATGTCCCTACT--GGGAATPACAGACTGACAGACTGGAGCATCTTAAAGTGC 961
Db      883 AACATGCCACTTACTGGGGAACAGATTAATCTTCCAGCTTTGGTGTGAATGATGCTC 942
Qy      962 AACGTTTTCTATGACTTTAGTAGAGATGAGAGAGAGATGATCTTAAAGCATGG 1021
Db      943 TTTTCTTCAAGCAT-----GGAAAGCATAT 971
Qy      1022 TGAAGAGATTAATGTTTTTATATCAACATCTTATATTAATTGATTCATTTAGTTACA 1081
Db      972 GGAAGAGACCAACAGTTGTTACACCCACCTGTGTCTTGAATTCATTTGAATTCGA 1031
Qy      1082 GTGGTGTGTGATGATTTTCTATCTTTTCCCTGACGTTACTTCAAGTAAACACA 1141
Db      1032 GGGGATCTAGTGAAGATTTCTATCTTTCCCTTAAAGCTCACTTCAAG3----- 1084
Qy      1142 AACTTTCATCAGCCATGATCTATAGAGACCTCTTAAGAGATATCGGTGATTTGG 1201
Db      1085 ---GTCCTTTTGTGACAGAGTACGGGCTGTCTTCAAGTCTGTCTGAGCAAT---- 1137
Qy      1202 ACCCAAACCATCTCTCAAAGCATTATATCCATCATCGCTGATGTTTAAATCAGC 1261
Db      1138 -CACAAGCATCTCTCAAAAACATTAATATACAGGCATGCTGTATGATTTTCA---- 1191
Qy      1262 AGAAGCATGTTTATGTTTGTACAAAAGAAATTTTATGGTGGGATGAGGTATAG 1321
Db      1192 -----CTGTCCGTGTTGTTTTCAATTGTATGTGAAGGGCTTGGGGTGGATTGA 1245
Qy      1322 ACCATGATGATGATCTTCAAGCTTATTAATPAAAGATCTTAAATGGGCGAGAGACT 1381
Db      1246 AGAATGCAAGATGCTCTGGGATGATTTCAATTAAGAGATCTTAAATGAGATGAGACT 1305
Qy      1382 GTGACAAAGACACCTTAATATATGGTGTGATGCTGAGTAGCAAAATCTTCTGAAGCCA 1441
Db      1306 AGCAAAATCACTCTGAAAAATGATTCAGGCTCAAGAGCAATCCCTGGAACACA 1365
Qy      1442 AACTCTTTT 1450
Db      1366 GACTCTTTT 1374

```

```

RESULT 10
AF132979 2440 bp mRNA linear ROD 05-AUG-1999
LOCUS Mus musculus activation-induced cytidine deaminase (AID) mRNA,
DEFINITION complete cds.
ACCESSION AF132979.1 GI:5305727
VERSION AF132979.1
KEYWORDS Mus musculus (house mouse)

```

```

ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2440)
Muraiatsu,M., Sankaranand,V.S., Anant,S., Sugai,M., Kinoshita,K.,
Davidson,N.O., and Honjo,T.
TITLE Specific expression of activation-induced cytidine deaminase (AID),
a novel member of the RNA-editing deaminase family in germinal
center B cells
JOURNAL J. Biol. Chem. 274 (26), 18470-18476 (1999)
MEDLINE 99303612
PUBMED 10373455
REFERENCE 2 (bases 1 to 2440)
AUTHORS Muraiatsu,M. and Honjo,T.
TITLE Direct Submision
JOURNAL Submitted (04-MAR-1999) Medical Chemistry, Kyoto University, Konoe
Yoshida Sakyo-Ku, Kyoto 606-8501, Japan
FEATURES
source
1..2440
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..2440
/gene="AID"
93..689
/note="APOBEC-1 homologue"
/codon_start=1
/product="activation-induced cytidine deaminase"
/protein_id="AA041793.1"
/db_xref="GI:5305728"
/translation="MDSLIMKQKPLVFNKRWAKGRHETLYCYVVKRDSATSCSL
DPGLRNRSGCHVELFLRYISDMDDPGCYVETWFTMSPCYDCARVIAEFLRNP
NLSIRPTALVCEDEKAEPELRHLHAGVIGIMTPKDYFCNNTVEVNERERFK
AMEGHENSRILRLRLPLIYEVDLRDARMTGP"

```

```

ORIGIN
Query Match 21.4%; Score 603.4; DB 10; Length 2440;
Best Local Similarity 69.4%; Pred. No. 6.4e-108;
Matches 1006; Conservative 0; Mismatches 351; Indels 92; Gaps 10;

```

```

Qy      4 GAACATCATTAATTAAGTGAATTTTCTGCTTGAAGCTTCAAGGAGCAAG-AAG 62
Db      16 GAACACCTTGTCTTGAAGCAAGCTTCTTGGCTTAAGACTTGAAGGAGTCAAGAAAG 75
Qy      63 AACTCTGGACACCATATGAGACAGCTCTTATGAAACCGAGAGATTTCTTTACCAAT 122
Db      76 TCAGCTGGAGACCGATATGAGACAGCTCTTATGAAACAAAGAAATTTCTTTACCAAT 135
Qy      123 TCAAAAATGTCGGTGGGCTTAAGGGTGGGCTGAGACTTACTGTGCTACGTAGTGAAG 182
Db      136 TCAAAAATGTCGGTGGGCTTAAGGGGCTGAGACTTACTGTGCTACGTAGTGAAG 195
Qy      183 GAGCTGACAGTGTACATCTTTTCACTGACCTTTGTTATCTTCCAAATPAAAGCGCT 242
Db      196 GAGAGATATGACACTCTGCTCACTGAGACTTGGCAACTTCCCAAAATGTCGGCT 255
Qy      243 GCAACGTGGAATGCTTTCTTCCGCTACATCTGAGCTGGAACCTTGAACCTTGGCGCT 302
Db      256 GCAACGTGGAATGTTGTTCTTCAAGCTTCACTCACTCACTGAGACCTTGAACCTTGGCGCT 315
Qy      303 GCTACGCGTCACTGCTGCTACCTCTGAGAGCCCTGCTACAGACTGAGCCGACATGTG 362
Db      316 GTTACGCGTCACTGCTGCTACCTCTGAGAGCCCTGCTACAGACTGAGCCGACATGTG 375
Qy      363 CCGACTTTCGAGAGGGAACCCCAACTCAGTGTGAGATCTTCAACCGGCGGCTTACT 422
Db      376 CTGAGTTTCTGAGATGGAACCTTACCTCAGCTGAGATTTTTCACGGCGGCTTACT 435
Qy      423 TCTGTGAGACCGGAAGGCTGAGCCGAGGGGCTGCGGGGCTGCAACCGCGCGGGGTGC 482
Db      436 TCTGTGAGACCGGAAGGCTGAGCCGAGGGGCTGCGGGGCTGCAACCGCGCGGGGTGC 495

```

```

QY 483 AAATAGCCATCATGACCTTCMAAGATTTTCTGCTGGAATCTTTGTGAAAC 542
Db 496 AGATCGGATCATACCTTCMAAGATTTTCTGCTGGAATCTTTGTGAAATC 555
QY 543 ATGAAAGAACTTTCAAGCTCGGAGAGGCTGATGAAATTCAGTCTCTCCAGAC 602
Db 556 GTGAAAGAACTTTCAAGCTCGGAGAGGCTGATGAAATTCAGTCTCTCCAGAC 615
QY 603 AGCTTCGGCGATCTTTTGGCCCTGTATGAGTTGATGCTTACGAGACGCTTTCGA 662
Db 616 AACTTCGGCGATCTTTTGGCCCTGTATGAGTTGATGCTTACGAGACGCTTTCGA 675
QY 663 CTTCGGGACTTTGATGAACTTCAGAAATGTCACACAGATGAAATATCTGCTGAA 722
Db 676 TGTTCGGATTTTGAAGCACTCTCTGGAATGTCACAGATGAAATATCTGCTGAA 732
QY 723 GACAGTGAATGAAAGAGCTCTCAAGCTCTCTGTTTATTTCTTCACTCTCACT 782
Db 733 GAGACTGATGAAAGAAACACCTTCA--ACTACATGTTTCTTCTTAACTGACTCACT 790
QY 783 TCTTAGATTTACGAAAAAATTTATATACGACTCTTTAAAGATCTTATGCTTGA 842
Db 791 TTATATGATGAGGGGAAA-----TTATATGACTTTTAAATACTTGAAGTGCAC 843
QY 843 AATAGAGAGGAAACAGAGTCTGCGCAGGAGCGTGCACATTTGTCAGTTTGAATGC 902
Db 844 AGAGCCCGCAGAGAAATGATGTATAC-----TGACCTTCGCTGTGC 882
QY 903 AACATTTGCTCCCTACT--GGAAATPACAGAACTGACAGACCTGGAGAGCTTAAGTGC 961
Db 883 AACATTCCTCATCTCTGGGAAACAGCATTAATCTTCAAGCTTTGGTGTGAATGATGCTC 942
QY 962 AACGTTTTCTATGACTTTAGTAGAGTGAAGCAGAGATGATCTTAAAGCATG 1021
Db 943 TTTTTCACACAGAT-----GGAAAGCATAT 971
QY 1022 TGAAGAGTCAAAAGTTTATATCAACATCTTATATTTGATTCATTTGATTAACA 1081
Db 972 GGAACAGACCAACAGATTTGTATCAACCCACCTGTGTTCTTATTCATTTGATTCACA 1031
QY 1082 GTGCTGTAGATGATGATTTTCTTATCTTCTTCCCTGAGCTTACTTCAAGTAAACACA 1141
Db 1032 GGGGTATCATGAGCGATTTCTTATTTCTTCTTCAAGCTCACTTTCAGG----- 1084
QY 1142 AACTCTTCATCAGGCGATGATCTATAGAGCTCTTATGAGATATCTGGTATGTTG 1201
Db 1085 ---GGTCTTTTTCAGAGAGGTACGGGGCTGTCTTCACTCTGTCTGAGCAAT----- 1137
QY 1202 ACCCCAAACATCTCTCCAAAGCATTAATATCCAAATCATGCGCTGATGTTTATCAGC 1261
Db 1138 -CAAAACCAATCTCTCAAAAACATTAATCTCAGGACATGCTGATGTTTCA----- 1191
QY 1262 AGAAGCATGTTTATGTTGTATCAAAAAGATGTTGATGGTGGGATGAGAGTATAG 1321
Db 1192 -----CTGTCGCGTGTGTTTTCATTTGTATGTAAGAGGGCTTGGGGTGGAGTTTGA 1245
QY 1322 ACCATGATGGTCACTTCAAGCTACTTATATAAGATCTTAAATAGGGCAGAGAGACT 1381
Db 1246 AGATGACGATGCGCTCTGCGGTATTTCAATTAAGATCTTAAATGACAGATGAGAGACT 1305
QY 1382 GTGAACAAGACACCTTATATATGAGTATGATCTGAAGTGAACAATCTTCTGAAACGCA 1441
Db 1306 ACAGAAGAAATCACTCTGAATATGATTCAGCGCTCAAGAGCAATCCCTGAAACACA 1365
QY 1442 AACCTTTT 1450
Db 1366 GACTCTTT 1374

```

RESULT 11
 LOCUS AF529828 597 bp mRNA linear SYN 17-SBP-2002
 DEFINITION Mus musculus clone 1 transgenic Homo sapiens AID (AID) mRNA,

```

ACCESSION complete cds.
VERSION AF529828
KEYWORDS AF529828.1 GI:22297243
SOURCE
ORGANISM Mus musculus (house mouse)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL 1 (bases 1 to 597)
TITLE Martin, A. and Schaff, M.D.
AUTHORS Somatic hypermutation of the AID transgene in B and non-B cells
REFERENCE Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
AUTHORS 2 (bases 1 to 597)
TITLE Martin, A. and Schaff, M.D.
AUTHORS Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
MEDICINE Medicine, 1300 Morris Park Ave. Chanin 404, Bronx, NY 10461, USA
FEATURES
SOURCE
1..597
location/qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="1"
/cell_type="hybridoma P1-5"
/transgenic
1..597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..597
/gene="AID"
1..597
/gene="AID"
/codon_start=1
/product="AID"
/protein_id="PAM5415.1"
/db_xref="GI:22297244"
/translation="MDSLNNRRKFLVQFNVRWAKGRRETYLCYVVRKDSATSFEL
DFGLRNNDCHVELFLARYISDWIDPGRCYRTWTSPSCYDCARHYADFGRNP
NLSRLIFARLYPCEDRKABEGRLRRHAGVOAIWTFPDYFCWNTPEHNHRTFK
AMBLHNSVRLSRLRIILPLVEVDLADAPRTTGL"
ORIGIN
Query Match 21.1%; Score 595.4; DB 12; Length 597;
Best Local Similarity 99.8%; Pred. No. 2.2e-106;
Matches 596; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 80 ATGACAGCCTCTTGATGACCGGAGAGATTTCTTTACCAATTCAAAAATGTCGCTGG 139
Db 1 ATGACAGCCTCTTGATGACCGGAGAGATTTCTTTACCAATTCAAAAATGTCGCTGG 60
QY 140 GCTAAGGCTGGCGTGAAGCTAAGTGTGCTAAGTGAAGAGCGGTGACGTCTACA 199
Db 61 GCTAAGGCTGGCGTGAAGCTAAGTGTGCTAAGTGAAGAGCGGTGACGTCTACA 120
QY 200 TCTTTTCACTGACCTTGTGTTATCTTTCGCAATTAAGAGCGTSCCAGTGAATTGCTC 259
Db 121 TCTTTTCACTGACCTTGTGTTATCTTTCGCAATTAAGAGCGTSCCAGTGAATTGCTC 180
QY 260 TTCCTCGCTACATCTCGGACTGGAACCTTAAGACCTTGGCGGTGCTACGCGTCACTGG 319
Db 181 TTCCTCGCTACATCTCGGACTGGAACCTTAAGACCTTGGCGGTGCTACGCGTCACTGG 240
QY 320 TTCACTCTCTGGAACCCCTGTGATGACTGTGCCCCGAGATGTGGCCGCACTTTTCGAGGG 379
Db 241 TTCACTCTCTGGAACCCCTGTGATGACTGTGCCCCGAGATGTGGCCGCACTTTTCGAGGG 300
QY 380 AACCCCAACCTCAGTGTGAGATCTTACACCGCGGCGCTTACTTCTGAGAGACCGCAAG 439
Db 301 AACCCCAACCTCAGTGTGAGATCTTACACCGCGGCGCTTACTTCTGAGAGACCGCAAG 360
QY 440 GCTAGCCGAGAGGGGCTGCGCGGCTGACCGCGCGGGGTGCAATATGACCATCATGACC 499

```

Db 361 GCTGAGCCGAGGCGCTGCGCGGCTGACCGCGCGGCGTGCAGAAATGACATCATGACC 420
QY 500 TTCAAGATATTTTATCTGCTGGAATACCTTTGTGAGAAACATGAAAGACTTTCAA 559
Db 421 TTCAAGATATTTTATCTGCTGGAATACCTTTGTGAGAAACATGAAAGACTTTCAA 480
QY 560 GCGTGGAGAGGCGTGCATGAAAATTCAGTTCGTCTCTCCAGACAGCTTGGCGCATCTT 619
Db 481 GCGTGGAGAGGCGTGCATGAAAATTCAGTTCGTCTCTCCAGACAGCTTGGCGCATCTT 540
QY 620 TTGCCCCCTGTATGAGTTGATGACTTACGAGACGACATTTGATCTTGGACATTTGA 676
Db 541 TTGCCCCCTGTATGAGTTGATGACTTACGAGACGACATTTGATCTTGGACATTTGA 597

RESULT 12
BT007402 597 bp mRNA linear PRI 13-MAY-2003
LOCUS Homo sapiens activation-induced cytidine deaminase mRNA, complete cds.
DEFINITION
ACCESSION BT007402
VERSION BT007402.1 GI:30583642
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Kaline, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S., Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y., Pheelan, M., and Farmer, A.
TITLE Cloning of human full-length CDS in BD Creator(TM) System Donor vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 597)
AUTHORS Kaline, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S., Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y., Pheelan, M., and Farmer, A.
TITLE Direct Submision
JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA

COMMENT This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'Acc' after SalI site and before 'Atc' to provide Kozak consensus sequence; 'Gg' after last codon and before HindIII site to maintain reading frame. Clone distribution: <http://Bioinfo.clontech.com/orfclones>. Location/Qualifiers
1..597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="BD Creator(TM) CDS Library derived from MGC collection"
/lab_host="DH5alpha T1 resistant"
/notes="vector: pDNR-Dual"
1..597
/codon_start=1
/product="activation-induced cytidine deaminase"
/protein_id="AAP3606.1"
/db_xref="GI:30583643"
/translation="MDSLIMNRKFLYQFKNVMAKGRRETYLCVVRKDSAFSFLDQYLNKNGCHVEHLFLRYISDWDIDPGRCYRVMTFSWSPCYCAHVAADPLRGNPNSIRIFATLYFCEDRKAPGEGRLRHAGVQIIMTFKDYFYCMNTEVENHERTFKAMEGHENSVRLSRQRLRLPLLYEVDLRDAFRLGL."

FEATURES
source
CDS
ORIGIN
Query Match 21.1%; Score 595; DB 9; Length 597;

Best Local Similarity 100.0%; Pred. No. 2 6e-106; Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 ATGACAGAGCTCTTGATGAAACCGAGAGAGTTTCTTTACCAATTCAAAAATGTCGCTGG 139
Db 1 ATGACAGAGCTCTTGATGAAACCGAGAGAGTTTCTTTACCAATTCAAAAATGTCGCTGG 60
QY 140 GCTAAGGTCGCGCTGAGACCTTACCTTGTCTAGTATGAAAGAGGCTGACAGTCTACA 199
Db 61 GCTAAGGTCGCGCTGAGACCTTACCTTGTCTAGTATGAAAGAGGCTGACAGTCTACA 120
QY 200 TCCTTTCACTGACCTTGGTTATCTTGCGAAATAGAAAGCGCTGCCACGNGAATTGCTC 259
Db 121 TCCTTTCACTGACCTTGGTTATCTTGCGAAATAGAAAGCGCTGCCACGNGAATTGCTC 180
QY 260 TTCTCCGCTACATCTCGAATCTGGAATGAGACCTTACCTTGGCTGACCTGACCTGG 319
Db 181 TTCTCCGCTACATCTCGAATCTGGAATGAGACCTTACCTTGGCTGACCTGACCTGG 240
QY 320 TTCACTTCTGAGACCTTCTGCTACGACTGTGCCAGATGTGCGCACTTTTGGCGAGG 379
Db 241 TTCACTTCTGAGACCTTCTGCTACGACTGTGCCAGATGTGCGCACTTTTGGCGAGG 300
QY 380 AACCCCAACCTCACTGAGGATCTTACCGCGGCGCTCACTTCTGTGAGACCGCAAG 439
Db 301 AACCCCAACCTCACTGAGGATCTTACCGCGGCGCTCACTTCTGTGAGACCGCAAG 360
QY 440 GCTGAGCCGAGGAGGCTGCGCGGCTGCAACCGCGCGGCTGCAAAATGACATCATGACC 499
Db 361 GCTGAGCCGAGGAGGCTGCGCGGCTGCAACCGCGCGGCTGCAAAATGACATCATGACC 420
QY 500 TTCAAGATATTTTATCTGCTGGAATACCTTTGTGAGAAACATGAAAGACTTTCAA 559
Db 421 TTCAAGATATTTTATCTGCTGGAATACCTTTGTGAGAAACATGAAAGACTTTCAA 480
QY 560 GCGTGGAGAGGCGTGCATGAAAATTCAGTTCGTCTCTCCAGACAGCTTGGCGCATCTT 619
Db 481 GCGTGGAGAGGCGTGCATGAAAATTCAGTTCGTCTCTCCAGACAGCTTGGCGCATCTT 540
QY 620 TTGCCCCCTGTATGAGTTGATGACTTACGAGACGACATTTGATCTTGGACATTTGA 674
Db 541 TTGCCCCCTGTATGAGTTGATGACTTACGAGACGACATTTGATCTTGGACATTTGA 595

RESULT 13
BT008226 597 bp mRNA SYN 13-MAY-2003
LOCUS Synthetic construct Homo sapiens activation-induced cytidine deaminase mRNA, partial cds.
DEFINITION
ACCESSION BT008226
VERSION BT008226.1 GI:30585290
KEYWORDS
SOURCE synthetic construct
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Kaline, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S., Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y., Pheelan, M., and Farmer, A.
TITLE Cloning of human full-length CDS in BD Creator(TM) System Donor vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 597)
AUTHORS Kaline, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S., Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y., Pheelan, M., and Farmer, A.
TITLE Direct Submision
JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal

tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES

SOURCE

```
1..597
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="GH000091.0"
/collection="BD Creator(TM) CDS library derived from MGC"
/lab_host="DH5alpha TI resistant"
/notes="Vector: pDNR-Dual"
1..>597
/notes="Mutations: 596:Stop->Leu"
/codon_start=1
/transl_table=1
/product="Homo sapiens activation-induced cytidine deaminase"
/protein_id="AAP6918.1"
/db_xref="GI:30585291"
/translation="MDSLIMRRKFLYQFNVMAGRETTYCTVVKRDSATFSLDFGLRNKGCHVELFLRYISDMDLPGRCVATWFTMSPCYDCAHYADFLRNP NLSIRITFARLYPCEDRKAPDEGRIRLRHAGVOIAIMTFDYCYCWNTPVENHERFFK AMEGLHENSVRLSRQLRRLILPLYEVDDLDAERITGL"
```

CDS

ORIGIN

```
Query Match      21.1%; Score 595; DB 12; Length 597;
Best Local Similarity 100.0%; Pred. No. 2.6e-106;
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 ATGACAGAGCTCTTGATGAACCGAGAGAGTTCTTACCAATTCAAAAATGCCCTGG 139
DB 1 ATGACAGAGCTCTTGATGAACCGAGAGAGTTCTTACCAATTCAAAAATGCCCTGG 60
QY 140 GCTAAGAGTGGCGCTGAGACCTACCTGTGCTAGTAGTAAGAGCGTGA CAGTGTCTACA 199
DB 61 GCTAAGAGTGGCGCTGAGACCTACCTGTGCTAGTAGTAAGAGCGTGA CAGTGTCTACA 120
QY 200 TCCTTTTCACTGACCTTGGTTATCTTGGCAATTAAGACGGCTGCCA CGTGAATTGCTC 259
DB 121 TCCTTTTCACTGACCTTGGTTATCTTGGCAATTAAGACGGCTGCCA CGTGAATTGCTC 180
QY 260 TTCTCTCGGCTACATCTGAGACTGAGACCTTGGCGCTGCTGCTACCG CGCTCACTGG 319
DB 181 TTCTCTCGGCTACATCTGAGACTGAGACCTTGGCGCTGCTGCTACCG CGCTCACTGG 240
QY 320 TTCACTCTGAGAGCGGCTGCTGACATGTGCCGCAATGTGGCCGACTT CTGGAGAGG 379
DB 241 TTCACTCTGAGAGCGGCTGCTGACATGTGCCGCAATGTGGCCGACTT CTGGAGAGG 300
QY 380 AACCCCAACCTCACTGAGAGATCTTCAACCGCGCGCTCTACTTCTGTG AGACCGCAAG 439
DB 301 AACCCCAACCTCACTGAGAGATCTTCAACCGCGCGCTCTACTTCTGTG AGACCGCAAG 360
QY 440 GCTGAGCCCGAGGGGCTGCGCGCGCTGACCGCGCGCGGGTGCATAA TATGCCATCATGACC 499
DB 361 GCTGAGCCCGAGGGGCTGCGCGCGCTGACCGCGCGCGGGTGCATAA TATGCCATCATGACC 420
QY 500 TTCAAGATATTTTCTGCTGTGGAATCTTTGTGTAAGAAACGATGAAG AACTTTTCAAA 559
DB 421 TTCAAGATATTTTCTGCTGTGGAATCTTTGTGTAAGAAACGATGAAG AACTTTTCAAA 480
QY 560 GCTGAGAGGGGCTGATGAATAATTCAGTTCCTCTCCAGACAGCTTGG CGCATCTT 619
DB 481 GCTGAGAGGGGCTGATGAATAATTCAGTTCCTCTCCAGACAGCTTGG CGCATCTT 540
QY 620 TTGCGCCTGTATGAGGTGATGACTTACGAGACGCAATTTGATCTTGG AACTTT 674
DB 541 TTGCGCCTGTATGAGGTGATGACTTACGAGACGCAATTTGATCTTGG AACTTT 595
```

```
RESULT 14
AF529815 596 bp mRNA linear PRI 17-SEP-2002
LOCUS Homo sapiens clone Ramos 1 AID (AID) mRNA, partial cds.
DEFINITION
ACCESSION AF529815
VERSION AF529815.1 GI:22297217
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
```

```
REFERENCE
AUTHORS Martin,A. and Scharff,M.D.
TITLE Somatic hypermutation of the AID transgene in B and non-B cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
REFERENCE
AUTHORS Martin,A. and Scharff,M.D.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave, Chanin 404, Bronx, NY 10461, USA
```

FEATURES

SOURCE

```
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 1"
1..>596
/gene="AID"
/gene="AID"
/notes="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="PAM95402.1"
/db_xref="GI:22297218"
/translation="MDSLIMRRKFLYQFNVMAGRETTYCTVVKRDSATFSLDFGLRNKGCHVELFLRYISDMDLPGRCVATWFTMSPCYDCAHYADFLRNP NLSIRITFARLYPCEDRKAPDEGRIRLRHAGVOIAIMTFDYCYCWNTPVENHERFFK AMEGLHENSVRLSRQLRRLILPLYEVDDLDAERITGL"
```

ORIGIN

```
Query Match      21.1%; Score 594.4; DB 9; Length 596;
Best Local Similarity 99.8%; Pred. No. 3.4e-106;
Matches 595; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 80 ATGACAGAGCTCTTGATGAACCGAGAGAGTTCTTACCAATTCAAAAATGCCCTGG 139
DB 1 ATGACAGAGCTCTTGATGAACCGAGAGAGTTCTTACCAATTCAAAAATGCCCTGG 60
QY 140 GCTAAGAGTGGCGCTGAGACCTACCTGTGCTAGTAGTAAGAGCGTGA CAGTGTCTACA 199
DB 61 GCTAAGAGTGGCGCTGAGACCTACCTGTGCTAGTAGTAAGAGCGTGA CAGTGTCTACA 120
QY 200 TCCTTTTCACTGACCTTGGTTATCTTGGCAATTAAGACGGCTGCCA CGTGAATTGCTC 259
DB 121 TCCTTTTCACTGACCTTGGTTATCTTGGCAATTAAGACGGCTGCCA CGTGAATTGCTC 180
QY 260 TTCTCTCGGCTACATCTGAGACTGAGACCTTGGCGCTGCTGCTACCG CGCTCACTGG 319
DB 181 TTCTCTCGGCTACATCTGAGACTGAGACCTTGGCGCTGCTGCTACCG CGCTCACTGG 240
QY 320 TTCACTCTGAGAGCGGCTGCTGACATGTGCCGCAATGTGGCCGACTT CTGGAGAGG 379
DB 241 TTCACTCTGAGAGCGGCTGCTGACATGTGCCGCAATGTGGCCGACTT CTGGAGAGG 300
QY 380 AACCCCAACCTCACTGAGAGATCTTCAACCGCGCGCTCTACTTCTGTG AGACCGCAAG 439
DB 301 AACCCCAACCTCACTGAGAGATCTTCAACCGCGCGCTCTACTTCTGTG AGACCGCAAG 360
QY 440 GCTGAGCCCGAGGGGCTGCGCGCGCTGACCGCGCGCGGGTGCATAA TATGCCATCATGACC 499
```

Db 361 GCTGAGCCGAGGGGCTGCGGCGCTGCACCGCGCGGGTGCATAATGACCATCATGACC 420
QY 500 TTCAAGATTAATTTTCTCTCTCTGGAATCTTTGTAGAAAACGATGAAGAACTTTCAA 559
Db 421 TTCAAGATTAATTTTCTCTCTGGAATCTTTGTAGAAAACGATGAAGAACTTTCAA 480
QY 560 GCTGGAAGGGGCTGCATGAAAATTCAGTTCTCTCTCCAGACAGCTTCGCGCATCTT 619
Db 481 GCTGGAAGGGGCTGCATGAAAATTCAGTTCTCTCTCCAGACAGATTCGCGCATCTT 540
QY 620 TTGCCCCGTATGAGGTGATGACTTACGAGACGCAATTTGTAATTGGGACTTTG 675
Db 541 TTGCCCCGTATGAGGTGATGACTTACGAGACGCAATTTGTAATTGGGACTTTG 596

RESULT 15
AF529816
LOCUS AF529816 596 bp mRNA linear PRI 17-SEP-2002
DEFINITION Homo sapiens clone Ramos 2 AID (AID) mRNA, partial cds.
ACCESSION AF529816
VERSION AF529816.1 GI:22297219
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 596)
Martin, A. and Scharff, M.D.
Somatic hypermutation of the AID transgene in B and non-B cells
Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
2 (bases 1 to 596)
Martin, A. and Scharff, M.D.
Direct Submision
Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Bronx, NY 10461, USA
FEATURES
source
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 2"
1..>596
/gene="AID"
1..>596
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/product="AID"
/codon_start=1
/protein_id="PAM95403.1"
/db_xref="GI:22297220"
/translation="MDSILMNRKFLYQFANVWAKRRETYLYVVKRSDTSFSL
DFGLRNKNGCHGELLFLRYISDMDLPGRCVYVFTSRSPCYDCARHADVFLRGNP
NLSRLFTARYFCEDRKAPRGLRIHAGVQIALMTFKDYFCMNTVENHETFK
AMRGLHENSVALSRQLRRLPLDYVDLDRDAFTLLGL"

ORIGIN
Query Match 21.1%; Score 594.4; DB 9; Length 596;
Best Local Similarity 99.8%; Pred. No. 3.4e-106;
Matches 595; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 80 ATGAGACGCTCTTGATGAACCGAGAGAGTTCTTACCAATTCAAAATGTCGCTG 139
Db 1 ATGAGACGCTCTTGATGAACCGAGAGAGTTCTTACCAATTCAAAATGTCGCTG 60
QY 140 GCTAAGGCTGCGCTGAGACCTTACCTGTCTAGTAGTAAGAGCGGTGACGTGTACA 199
Db 61 GCTAAGGCTGCGCTGAGACCTTACCTGTCTAGTAGTAAGAGCGGTGACGTGTACA 120
QY 200 TCCCTTTACCTGAGCTTTGGTTATCTTCGCAATPAGAAGCGGTGCAAGTGAATGCTC 259
Db 121 TCCCTTTACCTGAGCTTTGGTTATCTTCGCAATPAGAAGCGGTGCAAGTGAATGCTC 180
QY 260 TTCTCTCCGTACATCTCGAAGTGGAGCTGACCTGCGCTGTACCGCTCACTTG 319

Db 181 TTCTCCGCTACATCTCGAAGTGGAGACTTGAACCTTGGCCGCTGCTACCGGTTACCTTG 240
QY 320 TTCACTTCCTGAGACCCCTGCTACGACTGTGCGCGAAGTGTGCGCACTTTCTGCGAGGG 379
Db 241 TTCACTTCCTGAGACCCCTGCTACGACTGTGCGCGAAGTGTGCGCACTTTCTGCGAGGG 300
QY 380 AACCCCAACCTCACTTGAAGATTTTCAACCGCGGCTCTTACTTCTGTAGAACCGCAAG 439
Db 301 AACCCCAACCTCACTTGAAGATTTTCAACCGCGGCTCTTACTTCTGTAGAACCGCAAG 360
QY 440 GCTGAGCCGAGGGGCTGCGCGGCTGCAACCGCGCGGGTGCATAATGACATGACACC 499
Db 361 GCTGAGCCGAGGGGCTGCGCGGCTGCAACCGCGCGGGTGCATAATGACATGACACC 420
QY 500 TTCAAGATTAATTTTCTCTCTGGAATCTTTGTAGAAAACGATGAAGAACTTTCAA 559
Db 421 TTCAAGATTAATTTTCTCTCTGGAATCTTTGTAGAAAACGATGAAGAACTTTCAA 480
QY 560 GCTGGAAGGGGCTGCATGAAAATTCAGTTCTCTCTCCAGACAGCTTCGCGCATCTT 619
Db 481 GCTGGAAGGGGCTGCATGAAAATTCAGTTCTCTCTCCAGACAGCTTCGCGCATCTT 540
QY 620 TTGCCCCGTATGAGGTGATGACTTACGAGACGCAATTTGTAATTGGGACTTTG 675
Db 541 TTGCCCCGTATGAGGTGATGACTTACGAGACGCAATTTGTAATTGGGACTTTG 596

Search completed: March 7, 2004, 09:57:47
Job time : 10908 secs

PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
PT target for drug development for immune-related diseases including
PT allergies.

PS Claim 3; Page 135-139; 174pp; Japanese.

XX The present sequence encodes human activation-induced cytidine deaminase
CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
CC antidiabetic, ophthalmological, anti-HIV and dermatological activities,
CC and can be used in gene therapy. AID polynucleotides are useful in
CC methods for identifying drugs for the treatment of B cell associated
CC immune system disorders, immunodeficiency diseases and allergies, such as
CC immunoglobulin A (Iga) deficiency disease, Iga nephritis, gamma-
CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
CC drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
CC telangiectasia, common variable immunodeficiency disorder, MHC (major
CC histocompatibility class II deficiency disease, AIDS (auto
CC immunodeficiency syndrome), elevated IGE disorder, and Igc subclass
CC selection disorder. The DNA sequences encoding AID may be used for gene
CC therapy and the antibodies to the AID protein may be used for diagnosis
CC and treatment of these disorders

XX Sequence 2818 BP; 868 A; 548 C; 626 G; 776 T; 0 U; 0 Other;

Query Match 100.0%; Score 2818; DB 3; Length 2818;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AGAAGACCATCATTAATGTAAGTGAAGTCTTCTGCGCTGAGACTTGACGGAGGAGCAAGA 60
DB 1 AGAAGACCATCATTAATGTAAGTGAAGTCTTCTGCGCTGAGACTTGACGGAGGAGCAAGA 60
QY 61 AGACACTGAGACACCACTATGACAGCCTCTTGATGAACCGAGAGAGTTCTTTACCA 120
DB 61 AGACACTGAGACACCACTATGACAGCCTCTTGATGAACCGAGAGAGTTCTTTACCA 120
QY 121 ATTCAAAATGTCGCGTGGGCTTAAGGTCGCGCTGAGACCTTCTGCTAGCTAGTGA 180
DB 121 ATTCAAAATGTCGCGTGGGCTTAAGGTCGCGCTGAGACCTTCTGCTAGCTAGTGA 180
QY 181 GAGGCGGACAGTGTCAATCTCTTTCACCTGACCTTGTAATCTCGAATAAAGACG 240
DB 181 GAGGCGGACAGTGTCAATCTCTTTCACCTGACCTTGTAATCTCGAATAAAGACG 240
QY 241 CTGCGACGTGGAATGCTCTTCTCTCGCTACATCTCGACCTGAGACCTTGAACCTTGGCG 300
DB 241 CTGCGACGTGGAATGCTCTTCTCTCGCTACATCTCGACCTTGAACCTTGGCG 300
QY 301 CTGCTACCGGTCACCTGCTTCACTCTGAGACCCCTGCTAGACTGTGCCGACATGT 360
DB 301 CTGCTACCGGTCACCTGCTTCACTCTGAGACCCCTGCTAGACTGTGCCGACATGT 360
QY 361 GGGCGAATTTCTGCGAGGAGAACCCCAACCTCACTGAGATTTTACCGCGGCTCTTA 420
DB 361 GGGCGAATTTCTGCGAGGAGAACCCCAACCTCACTGAGATTTTACCGCGGCTCTTA 420
QY 421 CTCTGTGAGGACCGCAGGCTGAGCCCGAGGAGGCTGCGCGGCTGACCGCGCGGCT 480
DB 421 CTCTGTGAGGACCGCAGGCTGAGCCCGAGGAGGCTGCGCGGCTGACCGCGCGGCT 480
QY 481 GCAATATGACCATGACCTTCAAGATTAATTTTATCTGCTGGAATCTTTTGTAGAAA 540
DB 481 GCAATATGACCATGACCTTCAAGATTAATTTTATCTGCTGGAATCTTTTGTAGAAA 540
QY 541 CCATGAAAGAACTTTCAAGACCTGAGGAGGCTGCAATGAATTCAGTTCCTCCAG 600
DB 541 CCATGAAAGAACTTTCAAGACCTGAGGAGGCTGCAATGAATTCAGTTCCTCCAG 600
QY 601 ACAGCTTGCGGCGACCTTTTGGCCCTGTATGAGTGAATGATTCGAGAGGACATTTTG 660
DB 601 ACAGCTTGCGGCGACCTTTTGGCCCTGTATGAGTGAATGATTCGAGAGGACATTTTG 660
```

```
QY 661 TACTTGGGACCTTGTATGACAACTTCCAGAAATGTCAACACGATGAATATCTGCTG 720
DB 661 TACTTGGGACCTTGTATGACAACTTCCAGAAATGTCAACACGATGAATATCTGCTG 720
QY 721 AAGACAGTGAATTAATAAAGCTCCTTCAAGTCTTCTGTTTATTTCTTCAACTCTAC 780
DB 721 AAGACAGTGAATTAATAAAGCTCCTTCAAGTCTTCTGTTTATTTCTTCAACTCTAC 780
QY 781 TTTCTTGAAGTTTACAGAAAAATATTATATACGACTCTTTAAAAAGATCTATGCTTG 840
DB 781 TTTCTTGAAGTTTACAGAAAAATATTATATACGACTCTTTAAAAAGATCTATGCTTG 840
QY 841 AAAATAGAGAGAACACAGGCTGCGACAGGACGCTGCAATTTGCTGAGTTTGAAT 900
DB 841 AAAATAGAGAGAACACAGGCTGCGACAGGACGCTGCAATTTGCTGAGTTTGAAT 900
QY 901 GCAACATTTGCTCCTTACTGGAATTAACAGAACTGACAGACCTGAGACATCTTAAAGTG 960
DB 901 GCAACATTTGCTCCTTACTGGAATTAACAGAACTGACAGACCTGAGACATCTTAAAGTG 960
QY 961 CAACGTTTTCTATGACTTTTATGATGATGAGAGCAGAAAGTATCTTAAAAAGCATG 1020
DB 961 CAACGTTTTCTATGACTTTTATGATGATGAGAGCAGAAAGTATCTTAAAAAGCATG 1020
QY 1021 GTGAGAGATCAATGTTTTATATCAACATCTTTATTTATTTGATTCATTTGATTTAAC 1080
DB 1021 GTGAGAGATCAATGTTTTATATCAACATCTTTATTTATTTGATTCATTTGATTTAAC 1080
QY 1081 AGTGTGTATGATGATGATTTTCTATTTCTTTCCCTGAGCTTTCCTTCAAGTAAAC 1140
DB 1081 AGTGTGTATGATGATGATTTTCTATTTCTTTCCCTGAGCTTTCCTTCAAGTAAAC 1140
QY 1141 AAACCTTTCATCAAGGACATGATCTATAGACCTCTTATGAGAGATCTGAGTATGT 1200
DB 1141 AAACCTTTCATCAAGGACATGATCTATAGACCTCTTATGAGAGATCTGAGTATGT 1200
QY 1201 GACCCCAACCATCTCTCCCAAGATTAATCAATCAATGCGCTATGTTTAAATCAG 1260
DB 1201 GACCCCAACCATCTCTCCCAAGATTAATCAATCAATGCGCTATGTTTAAATCAG 1260
QY 1261 CAGAAAGATGTTTTATGTTTGTACAAAAGATGTTATGAGGAGGAGTGAAGTAA 1320
DB 1261 CAGAAAGATGTTTTATGTTTGTACAAAAGATGTTATGAGGAGGAGTGAAGTAA 1320
QY 1321 GACCATGATGATCACTTCAAGCTACTTAAATTAAGATCTTAAATGAGGACAGAGAC 1380
DB 1321 GACCATGATGATCACTTCAAGCTACTTAAATTAAGATCTTAAATGAGGACAGAGAC 1380
QY 1381 TGTGAACAGACACCTTAATATGAGTGTGATGTGAAGTGAATCTTCTGAAAACGC 1440
DB 1381 TGTGAACAGACACCTTAATATGAGTGTGATGTGAAGTGAATCTTCTGAAAACGC 1440
QY 1441 AAACCTTTTAAAGAAATCCCTAATTTAGAAAACCAACCAACTTCAATATATAATA 1500
DB 1441 AAACCTTTTAAAGAAATCCCTAATTTAGAAAACCAACCAACTTCAATATATAATA 1500
QY 1501 GCAAAACATTTGAAGAAATGCTTGAATGTTGGGAGAGAAATCTATTTGCTCTCGT 1560
DB 1501 GCAAAACATTTGAAGAAATGCTTGAATGTTGGGAGAGAAATCTATTTGCTCTCGT 1560
QY 1561 GGGTCTCTTCACTCAGAAATGCAATCAAGTCAAGGTTGCTACATTTGTATGTGT 1620
DB 1561 GGGTCTCTTCACTCAGAAATGCAATCAAGTCAAGGTTGCTACATTTGTATGTGT 1620
QY 1621 GATGCTTCTCCCAAGATTAATTAATTAATTAAGAGTGTGAACAAAACAGATATATA 1680
DB 1621 GATGCTTCTCCCAAGATTAATTAATTAATTAAGAGTGTGAACAAAACAGATATATA 1680
QY 1681 AGCTGCAACCGTGGACACGCTCATAGTCTGAGCTTGGAGAGGTTGAGAGGAGGA 1740
DB 1681 AGCTGCAACCGTGGACACGCTCATAGTCTGAGCTTGGAGAGGTTGAGAGGAGGA 1740
QY 1741 TGCGTTGAACAAGTGTTCAGGCGCAGCTGAGCAACATTAACAAATCTGTCTCAA 1800
```

Dp	1741	TGGCTTGAACAAGGTGTTCAAGGCGACGCTGGGCAACTTAAAGATCTGTCTCTGAA	1800
Qy	1801	AAAAAAAAAAAAAAAAAGAAAGAGAGAGGCGGGCGTGGTGCACGCTGTATCC	1860
Dp	1801	AAAAAAAAAAAAAAAAAGAAAGAGAGGCGGGCGTGGTGCACGCTGTATCC	1860
Qy	1861	CAGCACTTTGGGAGGCGGAGCGGGCGGATCACTGTGTGAGAGATTTGAGACCAAGCT	1920
Dp	1861	CAGCACTTTGGGAGGCGGAGCGGGCGGATCACTGTGTGAGAGATTTGAGACCAAGCT	1920
Qy	1921	GAGCAACATGGCCAAACCCTGTCTACTCAAAATGCAAAATTAGCAGAGCGTGTAGC	1980
Dp	1921	GAGCAACATGGCCAAACCCTGTCTACTCAAAATGCAAAATTAGCAGAGCGTGTAGC	1980
Qy	1991	AGGCACTCTGTATCCCACTACTTGGGAGGCTGAGGCGAGAAATCGCTTAACCCAGGA	2040
Dp	1991	AGGCACTCTGTATCCCACTACTTGGGAGGCTGAGGCGAGAAATCGCTTAACCCAGGA	2040
Qy	2041	GGTGGAGGTGCAGTAAGCTGAGATCGAGCCGTGGCACTCCAGCTGGGCGACAAGACA	2100
Dp	2041	GGTGGAGGTGCAGTAAGCTGAGATCGAGCCGTGGCACTCCAGCTGGGCGACAAGACA	2100
Qy	2101	AGACTCTGTCTCAGAAAAAAAAAAAAAAAAAGAGAGAGAGAAAGAAACAATATTG	2160
Dp	2101	AGACTCTGTCTCAGAAAAAAAAAAAAAAAAAGAGAGAGAGAAAGAAACAATATTG	2160
Qy	2161	GGAAGAAAGATGGGGAGCAATTGCAGAGAAATTGTCTTATCCAAACAAATGTAAAGA	2220
Dp	2161	GGAAGAAAGATGGGGAGCAATTGCAGAGAAATTGTCTTATCCAAACAAATGTAAAGA	2220
Qy	2221	GCCAAATAGGAGATCCCTATTTGTCTTTGGTGCTATTTGGCCCTTAACAACGTCTTT	2280
Dp	2221	GCCAAATAGGAGATCCCTATTTGTCTTTGGTGCTATTTGGCCCTTAACAACGTCTTT	2280
Qy	2281	GACAGTGAAGAAATATTCAAGATTAACCATATCCCTGTGCCGTTATTACTAGAACCTT	2340
Dp	2281	GACAGTGAAGAAATATTCAAGATTAACCATATCCCTGTGCCGTTATTACTAGAACCTT	2340
Qy	2341	TGCAATGAAGTGAAGCAATCCAGGAAAATTGAATGCACAACTGTCTTATTTTAATC	2400
Dp	2341	TGCAATGAAGTGAAGCAATCCAGGAAAATTGAATGCACAACTGTCTTATTTTAATC	2400
Qy	2401	TTATTTGATCAATTAAGTTGTAAAGGTAAAAATTTGTAACCTGCATGATTCATTATAT	2460
Dp	2401	TTATTTGATCAATTAAGTTGTAAAGGTAAAAATTTGTAACCTGCATGATTCATTATAT	2460
Qy	2461	TTATATTAATTTGGGCTCTAATGATTTTTTATTAATGATTTCCCTTCTGATATATGA	2520
Dp	2461	TTATATTAATTTGGGCTCTAATGATTTTTTATTAATGATTTCCCTTCTGATATATGA	2520
Qy	2521	AATGAGTCTCAAAAGCTTCATTAATTATTAACCTTAGAAATGATCTAATPAACAAGTAT	2580
Dp	2521	AATGAGTCTCAAAAGCTTCATTAATTATTAACCTTAGAAATGATCTAATPAACAAGTAT	2580
Qy	2581	GTAATTTGTAATTCAGATTAATTTGGTGTGCTAGAACCCATTTCTTGATTTTATTAAGT	2640
Dp	2581	GTAATTTGTAATTCAGATTAATTTGGTGTGCTAGAACCCATTTCTTGATTTTATTAAGT	2640
Qy	2641	TTTATGACAGCAAAATTTGCTTCTGGCTCACTTTCAATCACTTAATTAATGATTAAT	2700
Dp	2641	TTTATGACAGCAAAATTTGCTTCTGGCTCACTTTCAATCACTTAATTAATGATTAAT	2700
Qy	2701	TTTGGAAAGCTGTGAAGATAAATACCAATTAATTAATTAATTAATTAATTAAT	2760
Dp	2701	TTTGGAAAGCTGTGAAGATAAATACCAATTAATTAATTAATTAATTAATTAAT	2760
Qy	2761	TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2818
Dp	2761	TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2818

ID ABS73287
XX ABS73287 standard; DNA; 2791 BP.
AC ABS73287;
DT 04-DEC-2002 (first entry)
DE DNA encoding human translocation del(12p) protein #2.

KW Chromosome aberration; oncogenic fusion protein; cancer;
KW proliferative disease; cellular protein isoform; heat shock protein 90;
KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
KW acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
KW rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.

OS Homo sapiens.
XX
XX
PN WO200269900-A2.
XX
PD 12-SEP-2002.
XX
PF 01-MAR-2002; 2002MO-US006518.
XX
PR 01-MAR-2001; 2001US-0272751P.
XX
PA (CONF-) CONFORMA THERAPEUTICS CORP.
XX
PI Fritz LC, Burrows FU;
XX
DR WPT, 2002-698710/75.
XX
DR P-PsDB; ABG95083.

PT Treating genetically-defined disease associated with chromosomal
PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
PT diseases, involves administering an inhibitor of heat shock protein 90.

PS Disclosure; page 246-247; 38pp; English.
XX
XX
XX

The invention describes a method of treating genetically-defined disease associated with chromosomal aberrations yielding oncogenic fusion proteins (I), treating cancerous cells containing (I) in a heterogeneous cell population, treating proliferative diseases associated with mutant protein or cellular protein isoforms (II) dependent on heat shock protein (HSP)-90, or selectively treating cells expressing (II) involving administering HSP90-inhibitor. The method is useful for treating genetically-defined disease with chromosomal aberration yielding oncogenic fusion protein, treating cancerous cells containing fusion protein in heterogeneous cell population, treating proliferative disease (e.g. rheumatoid arthritis or cancer) associated with mutant protein or cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g. p53), or selectively treating cells expressing mutant protein or cellular protein isoform in a patient heterozygous for (II). The method is useful for treating a disease e.g. haematopoietic disorder such as T or B cell lymphoma, chronic myeloid leukaemia (CML), APL, ALL, ANL and CMML, or a disease characterised by a solid tumour such as papillary thyroid carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and synovial sarcoma. The method is also useful for treating viral infections. This represents the DNA sequence of a chromosome aberration

Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;

Query Match 99.0%; Score 2791; DB 6; Length 2791;
Best Local Similarity 100.0%; Prd. NO. 0;
Matches 2791; Conservative 0; Mismatches 0; Indels 0; Gaps 0

4 GAACCATTCAATTAAGTGAAGTAGATTTTTCGGCGCTTGAGACTTCAGGAGGCAGAAGA 63
Db 1 GAACCATCATTAATTAAGTAGAGTAGATTTTTCGGCGCTTGAGACTTCAGGAGGCAGAAGA 60
64 CACTGTGGACACCACTCTATNGACACGCTCTTGATGAAACCGAGGAAGTTTCTTAACAATT 123

Db 61 CACTCTGAGACCACTATGACAGCCTCTGATGAAACCGAGAAAGTTCTTACCAAT 120
Qy 124 CAAAATGTCGCGGAGCTAAGGCTCGGCGTGAAGACTACTGTGCTACTATGAAG 183
Db 121 CAAAATGTCGCGGAGCTAAGGCTCGGCGTGAAGACTACTGTGCTACTATGAAG 180
Qy 184 GCGTGAAGTGTACATCTCTTTCATGAGACTTGTGTTATCTTCGAAATGAAGACGGCTG 243
Db 181 GCGTGAAGTGTACATCTCTTTCATGAGACTTGTGTTATCTTCGAAATGAAGACGGCTG 240
Qy 244 CCACTGGAATGTGCTTCTCTCGCTAATCTCGAATGAGACTGAGCCTGAGCGCTG 303
Db 241 CCACTGGAATGTGCTTCTCTCGCTAATCTCGAATGAGACTGAGCCTGAGCGCTG 300
Qy 304 CTACCGGCTGACCTGCTTCACTCTGAGACCCCTGCTGACCTGCTGACATGTGC 363
Db 301 CTACCGGCTGACCTGCTTCACTCTGAGACCCCTGCTGACCTGCTGACATGTGC 360
Qy 364 CGACTTCTGAGAGGAAACCCCAACCTGATGAGGATCTTCAACGCGGCGCTCTACT 423
Db 361 CGACTTCTGAGAGGAAACCCCAACCTGATGAGGATCTTCAACGCGGCGCTCTACT 420
Qy 424 CTGTGAGAGCCGCAAGGCTGAGCCGAGGAGCTGCGGCGCTGACCGCGCGGCTGAG 483
Db 421 CTGTGAGAGCCGCAAGGCTGAGCCGAGGAGCTGCGGCGCTGACCGCGCGGCTGAG 480
Qy 484 AATGCAATGACCTTCAAGATTTTATCTGCTGAAATCTTTGTAGAAACCA 543
Db 481 AATGCAATGACCTTCAAGATTTTATCTGCTGAAATCTTTGTAGAAACCA 540
Qy 544 TGAAGAACTTCAAGACCTGAGAGGCTGAGTGAATAATTCAGTCTGCTCCAGAA 603
Db 541 TGAAGAACTTCAAGACCTGAGAGGCTGAGTGAATAATTCAGTCTGCTCCAGAA 600
Qy 604 GCTTCGCGGATCTTTTGGCCCTGTATGAGTTGATGACTTACGAGACGATTTGTGC 663
Db 601 GCTTCGCGGATCTTTTGGCCCTGTATGAGTTGATGACTTACGAGACGATTTGTGC 660
Qy 664 TTTGGACTTTGATGACACTTCCAGGAATGTCAACAGATGAATATCTGCTGAAG 723
Db 661 TTTGGACTTTGATGACACTTCCAGGAATGTCAACAGATGAATATCTGCTGAAG 720
Qy 724 ACACTGATTAATAAAACAGTCTTCAAGTCTTCTGTTTATTTCTCACTCTCT 783
Db 721 ACACTGATTAATAAAACAGTCTTCAAGTCTTCTGTTTATTTCTCACTCTCTCT 780
Qy 784 CTTAGACTTACAGAAAAATTTATATAGACTCTTTAAAAAGATCTATGCTGAAA 843
Db 781 CTTAGACTTACAGAAAAATTTATATAGACTCTTTAAAAAGATCTATGCTGAAA 840
Qy 844 ATAGAGAGGAAACAGGCTGAGCAGAGAGCTGCTGCAATGCTGAGATTTGAATGA 903
Db 841 ATAGAGAGGAAACAGGCTGAGCAGAGAGCTGCTGCAATGCTGAGATTTGAATGA 900
Qy 904 ACATTTGCTCCCTTATGAGAAATACAGAACTGCAAGACTTGGAGACTCTTAAAGTCTCA 963
Db 901 ACATTTGCTCCCTTATGAGAAATACAGAACTGCAAGACTTGGAGACTCTTAAAGTCTCA 960
Qy 964 CGTTTCTATGACTTTTATGATGAGAGAGAGATGAGATGCTTAAAGAGAGG 1023
Db 961 CGTTTCTATGACTTTTATGATGAGAGAGAGATGAGATGCTTAAAGAGAGG 1020
Qy 1024 AGAGATCAATGTTTATATCAACATCTTATTTATTTGATTTGATTTGAAGTCAAGT 1083
Db 1021 AGAGATCAATGTTTATATCAACATCTTATTTATTTGATTTGATTTGAAGTCAAGT 1080
Qy 1084 GGTGTATGATATGATTTTCTATCTTTTCCCTGAGCTTACCTTCAAGTAAACAA 1143
Db 1081 GGTGTATGATATGATTTTCTATCTTTTCCCTGAGCTTACCTTCAAGTAAACAA 1140
Qy 1144 CTCTTCATGAGGCAATGATCTATAGACTCTTCAATGAGAGATCTGGGTGATTTGTAC 1203

Db 1141 CTCTTCATGAGGCAATGATCTATAGACTCTTCAATGAGAGATCTGGGTGATTTGTAC 1200
Qy 1204 CCGAAACATCTCTCCAAAGATTAATTCGAATCAGCGCTGATGTTTAAATCAGAG 1263
Db 1201 CCGAAACATCTCTCCAAAGATTAATTCGAATCAGCGCTGATGTTTAAATCAGAG 1260
Qy 1264 AAGCATGTTTATGTTTGTACAAAGAAAGTGTATGAGTGGAGATGAGATATGAG 1323
Db 1261 AAGCATGTTTATGTTTGTACAAAGAAAGTGTATGAGTGGAGATGAGATATGAG 1320
Qy 1324 CATGATGTCACCTTCAAGCTACTTAAATGAAGATCTTAAATGAGCAGAGACTGT 1383
Db 1321 CATGATGTCACCTTCAAGCTACTTAAATGAAGATCTTAAATGAGCAGAGACTGT 1380
Qy 1384 GACCAAGCACCTCTAATATATGAGTGTATGAGATGAGCAATCTTCTGAAACGAAA 1443
Db 1381 GACCAAGCACCTCTAATATATGAGTGTATGAGATGAGCAATCTTCTGAAACGAAA 1440
Qy 1444 CTCTTTAAGAGAGCTCTAATTTAGAAACCCCAACCTTCAATGATTAATAGCA 1503
Db 1441 CTCTTTAAGAGAGCTCTAATTTAGAAACCCCAACCTTCAATGATTAATAGCA 1500
Qy 1504 AACATTTGAAAGAGTGTCTTGAATGTTGGAGAGAGAAATCTATGCTCTGAGG 1563
Db 1501 AACATTTGAAAGAGTGTCTTGAATGTTGGAGAGAGAAATCTATGCTCTGAGG 1560
Qy 1564 TCTCTCATCTCAGAAATGCAATCAGCTCAGGTTGCTACATTTTGTATGTGTAT 1623
Db 1561 TCTCTCATCTCAGAAATGCAATCAGCTCAGGTTGCTACATTTTGTATGTGTAT 1620
Qy 1624 GCTTCTCCAAAGATTAATTAATCTAATTAAGAGTGTGACAAAGAAATGATAAGC 1683
Db 1621 GCTTCTCCAAAGATTAATTAATCTAATTAAGAGTGTGACAAAGAAATGATAAGC 1680
Qy 1684 TGCAGACCGTGGCAACGCTCATAGTCTAGCTCTTGGAGGTTGAGAGGAGATGG 1743
Db 1681 TGCAGACCGTGGCAACGCTCATAGTCTAGCTCTTGGAGGTTGAGAGGAGATGG 1740
Qy 1744 CTTGAAACAGGTTTCAAGGCGACCTGAGCAATTAACAGATCTGTCTCAAAA 1803
Db 1741 CTTGAAACAGGTTTCAAGGCGACCTGAGCAATTAACAGATCTGTCTCAAAA 1800
Qy 1804 AAAAAAAAAAAAAAAAAAGAGAGAGGCGCGGCTGTGCTACGCTGTATTCAG 1863
Db 1801 AAAAAAAAAAAAAAAAAAGAGAGAGGCGCGGCTGTGCTACGCTGTATTCAG 1860
Qy 1864 CACTTTGAGAGCGGAGCGGCGGAGATCACTGTGCTCAGAGTTGAGACGAGCTGGC 1923
Db 1861 CACTTTGAGAGCGGAGCGGCGGAGATCACTGTGCTCAGAGTTGAGACGAGCTGGC 1920
Qy 1924 CAACATGCAAAACCCGCTGTATCTCAAAATGCAAAATTAAGCAGGCTGTAGCAG 1983
Db 1921 CAACATGCAAAACCCGCTGTATCTCAAAATGCAAAATTAAGCAGGCTGTAGCAG 1980
Qy 1984 CACTGTATCTCAGCTATCTTGGAGGCTGAGGAGAGATGCTTGAACCCAGAGAGT 2043
Db 1981 CACTGTATCTCAGCTATCTTGGAGGCTGAGGAGAGATGCTTGAACCCAGAGAGT 2040
Qy 2044 GAGGTTGCAATGAGCTGAGATGCTGCGGTTGCACTCAGCTGGGAGCAAGAGCAAG 2103
Db 2041 GAGGTTGCAATGAGCTGAGATGCTGCGGTTGCACTCAGCTGGGAGCAAGAGCAAG 2100
Qy 2104 CTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2163
Db 2101 CTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
Qy 2164 GAGAGAGTGGAG 2223
Db 2161 GAGAGAGTGGAG 2220
Qy 2224 AATTAAGGATCCCTAATTTGTCTCTTTGTGTCTAATTTGTCTTCAACATCTGCTTGA 2283
Db 2221 AATTAAGGATCCCTAATTTGTCTCTTTGTGTCTAATTTGTCTTCAACATCTGCTTGA 2280

QY 2284 AGTGAAGAAATATTCAGAAATACCATATCCCTGCGCTTATTAACCTAGCAACCCCTGG 2343
 Db 2281 AGTGAAGAAATATTCAGAAATACCATATCCCTGCGCTTATTAACCTAGCAACCCCTGG 2340
 QY 2344 AATGAATGAGCAGATCCACAGGAAACTGGAATGACAACTGCTTATTTTATCTTA 2403
 Db 2341 AATGAATGAGCAGATCCACAGGAAACTGGAATGACAACTGCTTATTTTATCTTA 2400
 QY 2404 TTGACATTAAGTTGTAAAGAGTAAAGTAAATGTCTTCACTGATTAATTAATTTA 2463
 Db 2401 TTGACATTAAGTTGTAAAGAGTAAAGTAAATGTCTTCACTGATTAATTAATTTA 2460
 QY 2464 TATATTTTGGCTTAAATGATTTTATTAACATGATTTCTTCTGATTAATTTGAAT 2523
 Db 2461 TATATTTTGGCTTAAATGATTTTATTAACATGATTTCTTCTGATTAATTTGAAT 2520
 QY 2524 GGAGCTCAAGCTTCAATTAATTTATTAATCTTGAATGAAATGCTTAATTAACAGCTATGTA 2583
 Db 2521 GGAGCTCAAGCTTCAATTAATTTATTAATCTTGAATGAAATGCTTAATTAACAGCTATGTA 2580
 QY 2584 ATTGTAACTTGCAGATATGCTGCTACAGGACCAATTTCTTGAATTTAGTAACTTT 2643
 Db 2581 ATTGTAACTTGCAGATATGCTGCTACAGGACCAATTTCTTGAATTTAGTAACTTT 2640
 QY 2644 ATGACAGCAATTTGCTTCTGCTCACTTCACTGATTAATTAATGATTAATTTT 2703
 Db 2641 ATGACAGCAATTTGCTTCTGCTCACTTCACTGATTAATTAATGATTAATTTT 2700
 QY 2704 GGAAGCTGTGAAGTAAATATACCAATTAATTAATTAAGATTTATGAAAGTTAA 2763
 Db 2701 GGAAGCTGTGAAGTAAATATACCAATTAATTAATTAAGATTTATGAAAGTTAA 2760
 QY 2764 AATTAATAATCAGTATGATGAAATTAACCTTG 2794
 Db 2761 AATTAATAATCAGTATGATGAAATTAACCTTG 2791

RESULT 3
 ABS73288
 ID ABS73288 standard; DNA; 2791 BP.
 XX
 AC ABS73288;

DT 04-DEC-2002 (first entry)

DE DNA encoding human translocation del(12p) protein #3.

XX Chromosome aberration; oncogenic fusion protein; cancer;
 KM Proliferative disease; cellular protein isoform; heat shock protein 90;
 KM HSP-90; Rheumatoid arthritis; cancer; haematopoietic disorder;
 KM T cell lymphoma; B cell lymphoma; chronic myeloid leukemia; CML;
 KM acute myeloid leukemia; AML; chronic myelomonocytic leukemia; CMML;
 KM papillary thyroid carcinoma; ALL; APL; NHL; solid tumour;
 KM papillary thyroid carcinoma; Ewing's sarcoma; melanoma; rhabdomyosarcoma;
 KM rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
 OS Homo sapiens.

XX MO200269500-A2.

XX 12-SEP-2002.

XX 01-MAR-2002; 2002MO-US006518.

XX 01-MAR-2001; 2001US-0272751P.

XX (CONF-) CONFORMA THERAPEUTICS CORP.

XX Filtz LC, Burows FU;

XX WPI, 2002-688710/75.

XX P-PSDB; AB95084.

XX Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 diseases, involves administering an inhibitor of heat shock protein 90.
 XX
 XX Disclosure; Page 248-249; 389pp; English.
 CC The invention describes a method of treating genetically-defined disease
 CC associated with chromosomal aberrations yielding oncogenic fusion
 CC proteins (I), treating cancerous cells containing (II) in a heterogeneous
 CC cell population, treating proliferative diseases associated with mutant
 CC protein or cellular protein isoforms (II) dependent on heat shock protein
 CC (HSP)-90, or selectively treating cells expressing (II) involving
 CC administering HSP90-inhibitor. The method is useful for treating
 CC genetically-defined disease with chromosomal aberration yielding
 CC oncogenic fusion protein, treating cancerous cells containing fusion
 CC protein in heterogeneous cell population, treating proliferative disease
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
 CC p53), or selectively treating cells expressing mutant protein or cellular
 CC protein isoform in a patient heterozygous for (II). The method is useful
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell
 CC lymphoma, chronic myeloid leukemia (CML), APL, ALL, NHL and CMML,
 CC or a disease characterised by a solid tumour such as papillary thyroid
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
 CC synovial sarcoma. The method is also useful for treating viral
 CC infections. This represents the DNA sequence of a chromosome aberration
 XX
 SQ Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;
 Query Match 99.0%; Score 2791; DB 6; Length 2791;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2791; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 GAACATCATTAATTAAGTGAAGTATTTTCTGCTGAGACTTGACAGAGGAGCAAGAGA 63
 Db 1 GAACATCATTAATTAAGTGAAGTATTTTCTGCTGAGACTTGACAGAGGAGCAAGAGA 60
 QY 64 CACTCTGACACACACTATGACAGCCTCTGATGAAACCGAGAGAGTTCTTTACCAATT 123
 Db 61 CACTCTGACACACACTATGACAGCCTCTGATGAAACCGAGAGAGTTCTTTACCAATT 120
 QY 124 CAATAATGTCGCTGAGCTAAGAGTGCAGTGAAGCCTACCTGTCTAGTGAAGAG 183
 Db 121 CAATAATGTCGCTGAGCTAAGAGTGCAGTGAAGCCTACCTGTCTAGTGAAGAG 180
 QY 184 GCGTGAACAGTCAATCCTTTTCACTGACGACTTTGTTATCTTGCGCAATGAACGCGCTG 243
 Db 181 GCGTGAACAGTCAATCCTTTTCACTGACGACTTTGTTATCTTGCGCAATGAACGCGCTG 240
 QY 244 CCAGTGAATTTGCTTCTTCCTGCGCTACATCTGAGACTGAGACCTGAGCCGCGCTG 303
 Db 241 CCAGTGAATTTGCTTCTTCCTGCGCTACATCTGAGACTGAGACCTGAGCCGCGCTG 300
 QY 304 CTACCGGCTCACTGTTTCACTCTGAGACCCCTGCTACAGCTGTGCCGCAATGTGC 363
 Db 301 CTACCGGCTCACTGTTTCACTCTGAGACCCCTGCTACAGCTGTGCCGCAATGTGC 360
 QY 364 CGACTTTCTGAGAGGAAACCCCAACTCACTGAGAGATTTCAACCGCGGCTTACTT 423
 Db 361 CGACTTTCTGAGAGGAAACCCCAACTCACTGAGAGATTTCAACCGCGGCTTACTT 420
 QY 424 CTGTGAAGACCGAAGGCTGAGCCGAGGCGCTGCGGCTGCAACCGCGCGGCTGCA 483
 Db 421 CTGTGAAGACCGAAGGCTGAGCCGAGGCGCTGCGGCTGCAACCGCGCGGCTGCA 480
 QY 484 AATAGCATCATGACCTTCAAGATTAATTTTCTGCTGAGAACTTTGTAGAAACCA 543
 Db 481 AATAGCATCATGACCTTCAAGATTAATTTTCTGCTGAGAACTTTGTAGAAACCA 540
 QY 544 TGAAGAACTTTCAAGCCTGGAAGGCTGATGAATTAATTCAGTTCTCTCCAGACA 603
 Db 541 TGAAGAACTTTCAAGCCTGGAAGGCTGATGAATTAATTCAGTTCTCTCCAGACA 600

QY 604 GCTTCGGCGCATCTTTTGGCCCTGTATGAGTGTGATGACCTTACGAGACGATTCGTGAC 663
Db 601 GCTTCGGCGCATCTTTTGGCCCTGTATGAGTGTGATGACCTTACGAGACGATTCGTGAC 660
QY 664 TTTGGGACTTTGATGACCACTTCAGGAATGTGACACAGATGAATATCTCTGTGAAG 723
Db 661 TTTGGGACTTTGATGACCACTTCAGGAATGTGACACAGATGAATATCTCTGTGAAG 720
QY 724 ACAGTGTATAAAAACAGTCTTCAAGTCTTCTGTGTTTATCTTCACTCTCACTT 783
Db 721 ACAGTGTATAAAAACAGTCTTCAAGTCTTCTGTGTTTATCTTCACTCTCACTT 780
QY 784 CTTAGAGTTTACGAAAAATTTTATATATGACCTTTTAAAAAGATCTATGCTTGAA 843
Db 781 CTTAGAGTTTACGAAAAATTTTATATATGACCTTTTAAAAAGATCTATGCTTGAA 840
QY 844 ATAGAGAGAGACACAGTCTGCGCAAGGACGTGTGCAATTTGTCATGTTTGAATGCA 903
Db 841 ATAGAGAGAGACACAGTCTGCGCAAGGACGTGTGCAATTTGTCATGTTTGAATGCA 900
QY 904 ACATGTCCCTTACTGGAATTAACAGACTGACAGACTTGGAAGCATCTTAAAGTGTCA 963
Db 901 ACATGTCCCTTACTGGAATTAACAGACTGACAGACTTGGAAGCATCTTAAAGTGTCA 960
QY 964 CGTTTTCTATGACTTTTAGTGAATGAGACAGAGAGATGATCTTAAAAAGCATGTG 1023
Db 961 CGTTTTCTATGACTTTTAGTGAATGAGAGAGAGATGATCTTAAAAAGCATGTG 1020
QY 1024 AGAGATCAATGTTTATATCAACATCTTATATTTGATTTCACTTGAATTAACAT 1083
Db 1021 AGAGATCAATGTTTATATCAACATCTTATATTTGATTTCACTTGAATTAACAT 1080
QY 1084 GGTGTAGTATGATTTTCTATCTTTCCCTGACGTTTACTTCAATGACACAA 1143
Db 1081 GGTGTAGTATGATTTTCTATCTTTCCCTGACGTTTACTTCAATGACACAA 1140
QY 1144 CTCTTCATCAAGCCATGATCTATAGGACCTCTTAATGAGATTTGGGTGATGTGAC 1203
Db 1141 CTCTTCATCAAGCCATGATCTATAGGACCTCTTAATGAGATTTGGGTGATGTGAC 1200
QY 1204 CCCAAACCATCTGCCAAAGCATTAATATCCATTCAGGCTGATGTTTATACAGAC 1263
Db 1201 CCCAAACCATCTGCCAAAGCATTAATATCCATTCAGGCTGATGTTTATACAGAC 1260
QY 1264 AAGCATGTTTATGTTTGTACAAAAGAGATTTTATGGGTGGGATGAGGTATAGAC 1323
Db 1261 AAGCATGTTTATGTTTGTACAAAAGAGATTTTATGGGTGGGATGAGGTATAGAC 1320
QY 1324 CATGATGTGACCTTCAAGCTTCTTATTAAGATCTTAAAAATGGGACAGAGACTGT 1383
Db 1321 CATGATGTGACCTTCAAGCTTCTTATTAAGATCTTAAAAATGGGACAGAGACTGT 1380
QY 1384 GAGCAAGACACCTTAATATAGGTTGATGTCTGAATTAAGCAATCTTCTGGAAGCGCAA 1443
Db 1381 GAGCAAGACACCTTAATATAGGTTGATGTCTGAATTAAGCAATCTTCTGGAAGCGCAA 1440
QY 1444 CTCTTTAAGGAAGTCCCTAATTTAGAAACACCAACCACTTCACTATCATATTAAGA 1503
Db 1441 CTCTTTAAGGAAGTCCCTAATTTAGAAACACCAACCACTTCACTATCATATTAAGA 1500
QY 1504 AACAAATGGAAGAGTGTGTTGATGTTGGGAGAGAAATCTATTTGCTCTGTGG 1563
Db 1501 AACAAATGGAAGAGTGTGTTGATGTTGGGAGAGAAATCTATTTGCTCTGTGG 1560
QY 1564 TCTCTTATCTCGAAATATGCAATCAAGTCAAGTTCCTCACTTTGTATGTGTAT 1623
Db 1561 TCTCTTATCTCGAAATATGCAATCAAGTCAAGTTCCTCACTTTGTATGTGTAT 1620
QY 1624 GCTTCTCCCAAGGTATTTACTATATTAAGAGTTGTGACAAACAGATGATTAAGC 1683
Db 1621 GCTTCTCCCAAGGTATTTACTATATTAAGAGTTGTGACAAACAGATGATTAAGC 1680

QY 1684 TGCAGCCGTGGCACAGCTCATAGTTCTAGCTGTTGGAGTTGAGAGGAGATGG 1743
Db 1681 TGCAGCCGTGGCACAGCTCATAGTTCTAGCTGTTGGAGTTGAGAGGAGATGG 1740
QY 1744 CTGAAACAGTGTTCAGGCCAGCCTGGGCAACATTAACAGATCTGTCTCTCAAAA 1803
Db 1741 CTGAAACAGTGTTCAGGCCAGCCTGGGCAACATTAACAGATCTGTCTCTCAAAA 1800
QY 1804 AAAAAAAAAAAAAAGAGAGAGGCCCGGGCGTGTGCTCAGGCCGTAAATCCAG 1863
Db 1801 AAAAAAAAAAAAAAGAGAGAGGCCCGGGCGTGTGCTCAGGCCGTAAATCCAG 1860
QY 1864 CACTTGGAGAGCCGAGCCGGGCGGATCACTGTGTCAGAGATTGAGACCAAGCTGGC 1923
Db 1861 CACTTGGAGAGCCGAGCCGGGCGGATCACTGTGTCAGAGATTGAGACCAAGCTGGC 1920
QY 1924 CAACATGGCAAAACCCGCTGTATCTCAAAATGCAAAATTAAGCAAGCGGTGTAGCAG 1983
Db 1921 CAACATGGCAAAACCCGCTGTATCTCAAAATGCAAAATTAAGCAAGCGGTGTAGCAG 1980
QY 1984 CACTGTATTCACAGCTTCTTTGGAGGCTGAGGACAGAGATGCTTGAACCAAGAGT 2043
Db 1981 CACTGTATTCACAGCTTCTTTGGAGGCTGAGGACAGAGATGCTTGAACCAAGAGT 2040
QY 2044 GGAGTTGCAAGTATGATGATGTCGCTGTGACTCCAGCCGTTGGGCAAGAGCAAG 2103
Db 2041 GGAGTTGCAAGTATGATGATGTCGCTGTGACTCCAGCCGTTGGGCAAGAGCAAG 2100
QY 2104 CTCTGTCTCAGAAAAAAGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2163
Db 2101 CTCTGTCTCAGAAAAAAGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
QY 2164 GAGAGAGATGGGAGAGATTCAGAGAAATTTGCTTTATCCAAACAAATGTAGAGCC 2223
Db 2161 GAGAGAGATGGGAGAGATTCAGAGAAATTTGCTTTATCCAAACAAATGTAGAGCC 2220
QY 2224 AATAGGATTCCTCATTTGTCTCTTTGGTGTCTATTTGCCCTTAACTGCTTTGAC 2283
Db 2221 AATAGGATTCCTCATTTGTCTCTTTGGTGTCTATTTGCCCTTAACTGCTTTGAC 2280
QY 2284 AGTGAAGAAATATTCAGATTAACATATCCCTGTGCGGTATTAACCTGCAACCTTGC 2343
Db 2281 AGTGAAGAAATATTCAGATTAACATATCCCTGTGCGGTATTAACCTGCAACCTTGC 2340
QY 2344 AATGAGATGACAGATCCACAGAAAACTTGAATGACACATGCTTATTTAATCTTA 2403
Db 2341 AATGAGATGACAGATCCACAGAAAACTTGAATGACACATGCTTATTTAATCTTA 2400
QY 2404 TTGTACATTAAGTTGTAAAGAGTTAAAAATTTGTACTGTATGATTTATATTTA 2463
Db 2401 TTGTACATTAAGTTGTAAAGAGTTAAAAATTTGTACTGTATGATTTATATTTA 2460
QY 2464 TATTAATTTGGTATATGATTTTATTAACATGATTTCTTCTGATATATGAAT 2523
Db 2461 TATTAATTTGGTATATGATTTTATTAACATGATTTCTTCTGATATATGAAT 2520
QY 2524 GAGTCTCAAGCTTCATTAATTTATTAACCTTGAAGATGTTCAATTAACAGATGTA 2583
Db 2521 GAGTCTCAAGCTTCATTAATTTATTAACCTTGAAGATGTTCAATTAACAGATGTA 2580
QY 2584 ATTGTAACTTGCAGATATAGTGTACAGAACCAATTTCTTGATTTTATGAATCTTT 2643
Db 2581 ATTGTAACTTGCAGATATAGTGTGTACAGAACCAATTTCTTGATTTTATGAATCTTT 2640
QY 2644 ATGACAGCAAAATTTGCTTGTGCTCACTTCAATCGTTAAATTAATGAATTAATTT 2703
Db 2641 ATGACAGCAAAATTTGCTTGTGCTCACTTCAATCGTTAAATTAATGAATTAATTT 2700
QY 2704 GGAAGCTGTGAAGATTAATACCAAAATTAATTAATTAATGAATTTATGAATTTA 2763
Db 2701 GGAAGCTGTGAAGATTAATACCAAAATTAATTAATTAATGAATTTATGAATTTA 2760
QY 2764 AATTAATAATCAATATGATGAATTAACCTTG 2794

OY	1697	CAAGCCTATAGTGTCTACTGCTGTGGAGAGTTGGAGAGAGAGATGCGCTTGAACAACAGT	1756
Db	4815	ACAGCGCTATAGTGTCTACTGCTGTGGAGAGTTGGAGAGAGAGATGCGCTTGAACAACAGT	4874
OY	1757	GTTCAAGGCCAGCCTGGGCAACATACACAGATCTGTCTCTCAAAAAAAAAAAAAAAAAAAAA	1816
Db	4875	GTTCAAGGCCAGCCTGGGCAACATACACAGATCTGTCTCTCAAAAAAAAAAAAAAAAAAAAA	4934
OY	1817	AAGAAAGAGAGAGGGCCGGGCGTGTGCTCAACGCTGTATTCGACACTTTGGAGGC	1876
Db	4935	AAGAAAGAGAGAGGGCCGGGCGTGTGCTCAACGCTGTATTCGACACTTTGGAGGC	4994
OY	1877	CGAGCGGGGGGAGATCACTGTGTGCTCAGGAGTTTGAGACACGCTGGGCAATGGCAAAA	1936
Db	4995	CGAGCGGGGGGAGATCACTGTGTGCTCAGGAGTTTGAGACACGCTGGGCAATGGCAAAA	5054
OY	1937	CCCCGTCTGTACTCAAAATGCAAAAATTAGCAGGCGGTGTAGAGAGCACTGTATATCC	1996
Db	5055	CCCCGTCTGTACTCAAAATGCAAAAATTAGCAGGCGGTGTAGAGAGCACTGTATATCC	5114
OY	1997	AGCTTACTTGGAGGCTGAGGCAAGAAATCGCTTGAACCCAGAGAGTGAAGTTGCAGTA	2056
Db	5115	AGCTTACTTGGAGGCTGAGGCAAGAAATCGCTTGAACCCAGAGAGTGAAGTTGCAGTA	5174
OY	2057	AGCTGAGATCGTCCGCTTGCACTCAGCCTGGGGGCAAGAGCAAGATCTGTCTTCAGAA	2116
Db	5175	AGCTGAGATCGTCCGCTTGCACTCAGCCTGGGGGCAAGAGCAAGATCTGTCTTCAGAA	5234
OY	2117	AAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGAAACAATTTGGGAGAGAGAGATGGG	2176
Db	5235	AAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGAAACAATTTGGGAGAGAGAGATGGG	5294
OY	2177	AAGCATTGCAAGAAATGTGCTTTATCCAACAAATGTGAAGACCAATPAAGGATCCC	2236
Db	5295	AAGCATTGCAAGAAATGTGCTTTATCCAACAAATGTGAAGACCAATPAAGGATCCC	5354
OY	2237	TATTTGTCTCTTTTGTGTCTATTTGTGCCCTPAACAACGTCTTTGACAGTGAAGAAATA	2296
Db	5355	TATTTGTCTCTTTTGTGTCTATTTGTGCCCTPAACAACGTCTTTGACAGTGAAGAAATA	5414
OY	2297	TTCAAAATPACATATCCCTGGCGGTATTAACGAGAACCCCTGCAATGAAGTAGC	2356
Db	5415	TTCAAAATPACATATCCCTGGCGGTATTAACGAGAACCCCTGCAATGAAGTAGC	5474
OY	2357	AGATCCACAGAGAAAATTGAATGCAACAACGTCTTATTTTATCTTATTTGTACATAGTT	2416
Db	5475	AGATCCACAGAGAAAATTGAATGCAACAACGTCTTATTTTATCTTATTTGTACATAGTT	5534
OY	2417	TGTAAAGAGTTAAAAATGTACTCTCAACGATATCATTTATATTTATATTTTGGCT	2476
Db	5535	TGTAAAGAGTTAAAAATGTACTCTCAACGATATCATTTATATTTATATTTTGGCT	5594
OY	2477	CTAATGATTTTTTATTAACGATTCCTTTTCTTTCGATATATTTGAAATGCAAGC	2536
Db	5595	CTAATGATTTTTTATTAACGATTCCTTTTCTTTCGATATATTTGAAATGCAAGC	5654
OY	2537	TTCAATAATTTATTAACCTTAGAAATGATCTTATTAACAACGATATGATTTGTAACATTC	2596
Db	5655	TTCAATAATTTATTAACCTTAGAAATGATCTTATTAACAACGATATGATTTGTAACATTC	5714
OY	2597	AGTAATGCTGCTAGAGAACCATTTCTCTGATTTTATAGTAACTTTATGACGAAATT	2656
Db	5715	AGTAATGCTGCTAGAGAACCATTTCTCTGATTTTATAGTAACTTTATGACGAAATT	5774
OY	2657	TGCTTCTGGCTCACCTTGAATCAGTTAATAATGATATTAATTTTGGAGCTGTGAG	2716
Db	5775	TGCTTCTGGCTCACCTTGAATCAGTTAATAATGATATTAATTTTGGAGCTGTGAG	5834
OY	2717	ATTAATAATCAAAATATAATATATATAAAGTGATTTATATGAAGTTAAATATAAATCAG	2776
Db	5835	ATTAATAATCAAAATATAATATATATAAAGTGATTTATATGAAGTTAAATATAAATCAG	5894
OY	2777	TATGATGGAATPAACCTGAAA	2797

Db 5895 TATGATGCAATTAACCTGAGA 5915

|||||

RESULT 5
AAC55339
ID AAC55339 standard; DNA; 11204 BP.
AAC55339;
05-FEB-2001 (first entry)

Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:35.

Activation-induced cytidine deaminase; AID; cytidine deaminase;
immune related disease; allergy; allergic disease; antiallergic;
antianemic; antisthmatic; ophthalmological; anti-HIV; dermatological;
gene therapy; B cell associated immune system disorder; food allergy;
immunodeficiency disease; immunoglobulin A deficiency disease; asthma;
IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
drug allergy; allergic rhinitis; Rosen disease; DiGeorge disease; AIDS;
ataxia telangiectasia; common variable immunodeficiency disorder;
major histocompatibility class II deficiency disease;
auto immunodeficiency syndrome; Igg subclass selection disorder; ds.

Homo sapiens.
WO200058480-A1.
05-OCT-2000.
28-MAR-2000; 2000MO-IP001918.
29-MAR-1999; 99JP-00087192.
24-JUN-1999; 99JP-00178999.
27-DEC-1999; 99JP-00371382.
(NIBS) JAPAN TOBACCO INC.
(HONJ/) HONJO T.
Honjo T, Muramatsu M;
WPI; 2000-611715/58.
Nucleic acid encoding activation induced cytidine deaminase, useful as a
target for drug development for immune-related diseases including
allergies.

Claim 17, Page 163-170, 174pp; Japanese.

The present invention describes an activation-induced cytidine deaminase
(AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
antisthmatic, ophthalmological, anti-HIV and dermatological activities,
and can be used in gene therapy. AID polynucleotides are useful in
methods for identifying drugs for the treatment of B cell associated
immune system disorders, immunodeficiency diseases and allergies, such as
immunoglobulin A (IGA) deficiency disease, Igg nephritis, gamma-
globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia
telangiectasia, common variable immunodeficiency disorder, MHC (major
histocompatibility class) II deficiency disease, AIDS (auto
immunodeficiency syndrome), elevated Igg disorder, and Igg subclass
selection disorder. The DNA sequences encoding AID may be used for gene
therapy and the antibodies to the AID protein may be used for diagnosis
and treatment of these disorders. The present sequence represents a
genomic DNA sequence of human AID

Sequence 11204 BP; 3305 A; 2273 C; 2373 G; 3253 T; 0 U; 0 Other;

Query Match 77.2%; Score 2174.6; DB 3; Length 11204;
Best Local Similarity 99.8%; Fred. No. 0;
Matches 2177; Conservative 0; Mismatches 4; Indels 0; Gaps 0

QY 617 CTTTGGCCCTGATGAGGTGATGACTTACGAGAGCATTTGCTGCTTGGACCTTGA 676
Db 8950 CTACAGCCCTGATGAGGTGATGACTTACGAGAGCATTTGCTGCTTGGACCTTGA 9009
QY 677 TAGCAACTTCAGGAATGTCAACAAGATGAATATCTGCTGGAAGACAGTGGATMAA 736
Db 9010 TAGCAACTTCAGGAATGTCAACAAGATGAATATCTGCTGGAAGACAGTGGATMAA 9069
QY 737 AACAGCTTCAGATCTCTCTGCTTTTATCTTCAACTCTCACTTTTAAGATTAA 796
Db 9070 AACAGCTTCAGATCTCTCTGCTTTTATCTTCAACTCTCACTTTTAAGATTAA 9129
QY 797 GAAAAAATTTATATACAGCTCTTAAAAAGATCTATGCTTGAATAAGAGAGAAC 856
Db 9130 GAAAAAATTTATATACAGCTCTTAAAAAGATCTATGCTTGAATAAGAGAGAAC 9189
QY 857 ACAAGTCTGGCCAGGACGTGCTGCAATTGTCAGATTTGATGACATTTGCCCTA 916
Db 9190 ACAAGTCTGGCCAGGACGTGCTGCAATTGTCAGATTTGATGACATTTGCCCTA 9249
QY 917 CTGGGAATTAACAACCTGACGACCTGGGAGACCTCTAAAGTGTCAACGTTTTCTAGA 976
Db 9250 CTGGGAATTAACAACCTGACGACCTGGGAGACCTCTAAAGTGTCAACGTTTTCTAGA 9309
QY 977 CTTTAAAGTATGAGAGAGAGAGAGATCTCTAAAGATGATGATGATGATGATGATG 1036
Db 9310 CTTTAAAGTATGAGAGAGAGAGAGATCTCTAAAGATGATGATGATGATGATGATG 9369
QY 1037 TTTTATATCAACATCTTTATATATTTGATCTATTGATGATTAACAGTGTGTTAGATA 1096
Db 9370 TTTTATATCAACATCTTTATATATTTGATCTATTGATGATTAACAGTGTGTTAGATA 9429
QY 1097 GATTTTCTATTTCTTCCCTGACGTTTACTTCAAGTACACAACCTCTTCCATCAG 1156
Db 9430 GATTTTCTATTTCTTCCCTGACGTTTACTTCAAGTACACAACCTCTTCCATCAG 9489
QY 1157 CCATGATCTATAGAGACCTCTCTATGAGATGATGATGATGATGATGATGATGATG 1216
Db 9490 CCATGATCTATAGAGACCTCTCTATGAGATGATGATGATGATGATGATGATGATG 9549
QY 1217 TCCAAAGCATTAATATCAATCATGCGCTGATGTTTAAATACAGAGAGACATGTTTAA 1276
Db 9550 TCCAAAGCATTAATATCAATCATGCGCTGATGTTTAAATACAGAGAGACATGTTTAA 9609
QY 1277 TGTGTTACAAAAGAGATGTTATGAGTGGGATGAGATGATGATGATGATGATGATG 1336
Db 9610 TGTGTTACAAAAGAGATGTTATGAGTGGGATGAGATGATGATGATGATGATGATGATG 9669
QY 1337 CTTCAAGCTACTTTAATAAGATCTTAAATGGGAGAGAGATGATGATGATGATGATG 1396
Db 9670 CTTCAAGCTACTTTAATAAGATCTTAAATGGGAGAGAGATGATGATGATGATGATG 9729
QY 1397 TAAATATGGGTGATGATCTGAGATGACAAATCTTCTGAAAGCAAATCTTTTAAGAA 1456
Db 9730 TAAATATGGGTGATGATCTGAGATGACAAATCTTCTGAAAGCAAATCTTTTAAGAA 9789
QY 1457 GTCCCTAATTTTGAACACCAACAACCTCACATATCAATTAAGCAAAATGGAAG 1516
Db 9790 GTCCCTAATTTTGAACACCAACAACCTCACATATCAATTAAGCAAAATGGAAG 9849
QY 1517 AAGTGTCTGAAATGTTGGGAGAGAGAAATCTATTTGCTCTGAGGCTCTCTTCACTCA 1576
Db 9850 AAGTGTCTGAAATGTTGGGAGAGAGAAATCTATTTGCTCTGAGGCTCTCTTCACTCA 9909
QY 1577 GAAATGCCATCAGGTCAGAGTGTGCTACATTTTGTATGATGATGATGATGATGATG 1636
Db 9910 GAAATGCCATCAGGTCAGAGTGTGCTACATTTTGTATGATGATGATGATGATGATG 9969
QY 1637 GTAATTAATCTATATAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1696
Db 9970 GTAATTAATCTATATAGAGATGATGATGATGATGATGATGATGATGATGATGATG 10029

QY 1697 AACGCTCATATGTTCTAGCTGCTTGGAGGTTGAGAGAGAGATGCTTGAACACAGGT 1756
Db 10030 AACGCTCATATGTTCTAGCTGCTTGGAGGTTGAGAGAGAGATGCTTGAACACAGGT 10089
QY 1757 GTTCAAGGCGACCTGGGCAACATTAACAAGATCTGCTCTCAAAAAAAAAAAAAAAAAA 1816
Db 10090 GTTCAAGGCGACCTGGGCAACATTAACAAGATCTGCTCTCAAAAAAAAAAAAAAAAAA 10149
QY 1817 AAGAAAGAGAGAGGCGGCGGTGTGTGCTCAAGCTGTATATCCAGACCTTTGGAGGC 1876
Db 10150 AAGAAAGAGAGAGGCGGCGGTGTGTGCTCAAGCTGTATATCCAGACCTTTGGAGGC 10209
QY 1877 CGAGCGGGGCGGATCACTGTGTGAGAGTTTGAACAGGCTGGGCAACATGGCAAAA 1936
Db 10210 CGAGCGGGGCGGATCACTGTGTGAGAGTTTGAACAGGCTGGGCAACATGGCAAAA 10269
QY 1937 CCCCCTGTACTCAAAATGCAAAAAATTAACAGGCGGTGTAGAGGCACTGTATATCCC 1996
Db 10270 CCCCCTGTACTCAAAATGCAAAAAATTAACAGGCGGTGTAGAGGCACTGTATATCCC 10329
QY 1997 AGCTACTTGGAGGCTGAGGCAAGAGATGCTTGAACCCAGAGGTTGACAGTA 2056
Db 10330 AGCTACTTGGAGGCTGAGGCAAGAGATGCTTGAACCCAGAGGTTGACAGTA 10389
QY 2057 AGCTGAGATGTCGCTGCTTGCATCTCAGCTGGGCGAACAAGCAACTGTCTCAGAA 2116
Db 10390 AGCTGAGATGTCGCTGCTTGCATCTCAGCTGGGCGAACAAGCAACTGTCTCAGAA 10449
QY 2117 AAAAAAAAAAAAAAAAAAG 2176
Db 10450 AAAAAAAAAAAAAAAAAAG 10509
QY 2177 AAGCATGCAAGAAATTTGCTTTATCCAACAAAATGTAAAGAGCCAAATAGAGATCCC 2236
Db 10510 AAGCATGCAAGAAATTTGCTTTATCCAACAAAATGTAAAGAGCCAAATAGAGATCCC 10569
QY 2237 TATTTGCTCTTTTGGTGTATTTTGTCCCTAACAACTGCTTTGACAGTGAAGAAAAA 2296
Db 10570 TATTTGCTCTTTTGGTGTATTTTGTCCCTAACAACTGCTTTGACAGTGAAGAAAAA 10629
QY 2297 TTCAAGATACCAATCCCTGTCCTGCTTATACCTAGCAACCTTGCATAGAAAGAGAG 2356
Db 10630 TTCAAGATACCAATCCCTGTCCTGCTTATACCTAGCAACCTTGCATAGAAAGAGAG 10689
QY 2357 AGATCCACAGAAAACCTGATGACAACTGTCTTATTTAATCTTATGATAGTAAAT 2416
Db 10690 AGATCCACAGAAAACCTGATGACAACTGTCTTATTTAATCTTATGATAGTAAAT 10749
QY 2417 TGTAAAGATTAATAATTTGTTAATTCATGATATTAATTTATATTTATTTTGGCT 2476
Db 10750 TGTAAAGATTAATAATTTGTTAATTCATGATATTAATTTATATTTATTTTGGCT 10809
QY 2477 CTATGATTTTTTATTAACATGATTTCTTTCTGATATTTGAATGAGAGCTCAAAAG 2536
Db 10810 CTATGATTTTTTATTAACATGATTTCTTTCTGATATTTGAATGAGAGCTCAAAAG 10869
QY 2537 TTCAATAATTTAATCTTAAAGAAATGATCTAATAACAAGTATGATTTGTAACATTCG 2596
Db 10870 TTCAATAATTTAATCTTAAAGAAATGATCTAATAACAAGTATGTAACATTCG 10929
QY 2597 AGTAATGTCATGAGAGCATTTCTCTGATTTTATGAACTTTATGACAGCAAT 2656
Db 10930 AGTAATGTCATGAGAGCATTTCTCTGATTTTATGAACTTTATGACAGCAAT 10989
QY 2657 TGTCTTGCTCACTTTCAATCAGTTAAATATATATATTTTGAAGCTGTGAAG 2716
Db 10990 TGTCTTGCTCACTTTCAATCAGTTAAATATATATATTTTGAAGCTGTGAAG 11049
QY 2717 ATAAATATCAATTAATAATTAATAATGATTTATGAGAGTTAAATTAATAATCG 2776
Db 11050 ATAAATATCAATTAATAATTAATAATGATTTATGAGAGTTAAATTAATAATCG 11109
QY 2777 TATGATGAGATAACTTGAAA 2797

Db 10030 ACAGCTCATAAGTTAGTGTCTGGAGAGTTGAGAGGAGAGATGGCTTGAACAGCT 10089
 QY 1757 GTTCAAGGCCAGCTGGGCACTAACAGATCTCTCTCAAAAAAAAAAAAAAAAAAAAA 1816
 Db 10090 GTTCAAGGCCAGCTGGGCACTAACAGATCTCTCTCAAAAAAAAAAAAAAAAAAAAA 10149
 QY 1817 AAGAAAGAGAGAGGCGCGGCGTGTGTGCTCAGCGCTGATCCGACGCTTGGAGGCG 1876
 Db 10150 AAGAAAGAGAGAGGCGCGGCGTGTGTGCTCAGCGCTGATCCGACGCTTGGAGGCG 10209
 QY 1877 CGAGCCGGGCGGATCACTGTGTGCTCAGAGTTGAGACGAGCTGGCCAACTGCAAAA 1936
 Db 10210 CGAGCCGGGCGGATCACTGTGTGCTCAGAGTTGAGACGAGCTGGCCAACTGCAAAA 10269
 QY 1937 CCCCCTGTGTACTCAAAATGCAAAATTTAGCCAGCGGTGTGAGGACGCTGTATGCC 1996
 Db 10270 CCCCCTGTGTACTCAAAATGCAAAATTTAGCCAGCGGTGTGAGGACGCTGTATGCC 10329
 QY 1997 AGCTACTGGGAGGCTGAGGACGAGAGATGCTTGAACCCAGAGGTGAGGTTGACAT 2056
 Db 10330 AGCTACTGGGAGGCTGAGGACGAGAGATGCTTGAACCCAGAGGTGAGGTTGACAT 10389
 QY 2057 AGCTGAGATGCTGCGCTTGCACTCCAGCTGGGCGAACAAGACAGACTGTGTGAGAA 2116
 Db 10390 AGCTGAGATGCTGCGCTTGCACTCCAGCTGGGCGAACAAGACAGACTGTGTGAGAA 10449
 QY 2117 AAAAAAAAAAAAAAAAAAG 2176
 Db 10450 AAAAAAAAAAAAAAAAAAG 10509
 QY 2177 AAGCTTGCAGAGAAATTTGCTTTATCCAACAAAATGTPAGAGCCAAATAGGATCCC 2236
 Db 10510 AAGCTTGCAGAGAAATTTGCTTTATCCAACAAAATGTPAGAGCCAAATAGGATCCC 10569
 QY 2237 TATTTGCTCTTTTGGTGTCTATTTGGTCCCTAACACTGTCTTTGCACTGAGAAAAA 2296
 Db 10570 TATTTGCTCTTTTGGTGTCTATTTGGTCCCTAACACTGTCTTTGCACTGAGAAAAA 10629
 QY 2297 TTGAGATTAACATATCCCTGTGCGCTTATTTACCTGACCAACCTTGCAATGAGAGC 2356
 Db 10630 TTGAGATTAACATATCCCTGTGCGCTTATTTACCTGACCAACCTTGCAATGAGAGC 10689
 QY 2357 AGATCCACAGAGAAACTTGAATGACACACTGTCTTATTTATCTTATTTGATAGT 2416
 Db 10690 AGATCCACAGAGAAACTTGAATGACACACTGTCTTATTTATCTTATTTGATAGT 10749
 QY 2417 TGTAAAAAGATTTAAATTTGTACTTCAATGATTTCAATTTATTTATTTTGGCT 2476
 Db 10750 TGTAAAAAGATTTAAATTTGTACTTCAATGATTTCAATTTATTTATTTTGGCT 10809
 QY 2477 CTAATGATTTTATTAACATGATTTCTTCTGATATGAAATGAGAGCTCAAGC 2536
 Db 10810 CTAATGATTTTATTAACATGATTTCTTCTGATATGAAATGAGAGCTCAAGC 10869
 QY 2537 TTCAATTAATTTATTAATTTGAATGATTTCTAATAACAAGTATGTAATTTGTAACCTTC 2596
 Db 10870 TTCAATTAATTTATTAATTTGAATGATTTCTAATAACAAGTATGTAATTTGTAACCTTC 10929
 QY 2597 AGTAATGCTCTACGAAGCAATTTCTTGTATTTTATTTAGTAATTTTATGACGCAAT 2656
 Db 10930 AGTAATGCTCTACGAAGCAATTTCTTGTATTTTATTTAGTAATTTTATGACGCAAT 10989
 QY 2657 TGCTCTGGCTCACTTCAATCAAGTAAATTAATTAATTAATTTTGAAGCTGTGAAG 2716
 Db 10990 TGCTCTGGCTCACTTCAATCAAGTAAATTAATTAATTAATTTTGAAGCTGTGAAG 11049
 QY 2717 ATAAATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2776
 Db 11050 ATAAATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11109
 QY 2777 TATGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2797

Db 11110 TATGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11130
 RESULT 7
 AAC55319
 ID AAC55319 standard; DNA; 2172 BP.
 XX AAC55319;
 AC
 XX 05-FEB-2001 (first entry)
 DT
 XX Human activation-induced cytidine deaminase exon 5 SEQ ID NO:15.
 DE
 XX
 XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
 KW immune related disease; allergy; allergic disease; anti-allergic;
 KW antianemic; antiaesthetic; ophthalmological; anti-HIV; dermatological;
 KW gene therapy; B cell associated immune system disorder; food allergy;
 KW immunodeficiency disease; immunoglobulin A deficiency disease; asplenia;
 KW IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;
 KW drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS;
 KW ataxia telangiectasia; common variable immunodeficiency disorder;
 KW major histocompatibility class II deficiency disease;
 KW auto immunodeficiency syndrome; IgG subclass selection disorder; ds.
 KW
 CS Homo sapiens.
 XX
 PN W0200058480-A1.
 XX
 PD 05-OCT-2000.
 XX
 PF 28-MAR-2000; 2000WC-JP001918.
 XX
 XX 29-MAR-1999; 99JP-00087192.
 PR 24-JUN-1999; 99JP-00178999.
 PR 27-DEC-1999; 99JP-00371382.
 XX
 PA (NISHI) JAPAN TOBACCO INC.
 PA (HONJ) HONJO T.
 PI Honjo T, Muramatsu M;
 XX
 DR WPI; 2000-611715/58.
 XX
 PT Nucleic acid encoding activation induced cytidine deaminase, useful as a
 PT target for drug development for immune-related diseases including
 PT allergies.
 XX
 PS Claim 18; Page 152-153; 174pp; Japanese.
 XX
 CC The present invention describes an activation-induced cytidine deaminase
 CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
 CC cytidine activity similar to APOBEC-1. AID has anti-allergic, antianemic,
 CC antiaesthetic, ophthalmological, anti-HIV and dermatological activities,
 CC and can be used in gene therapy. AID polynucleotides are useful in
 CC methods for identifying drugs for the treatment of B cell associated
 CC immune system disorder, immunodeficiency diseases and allergies, such as
 CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
 CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,
 CC drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia
 CC telangiectasia, common variable immunodeficiency disorder, MHC (major
 CC histocompatibility class II deficiency disease, AIDS (auto
 CC immunodeficiency syndrome), elevated IGE disorder, and IGG subclass
 CC selection disorder. The DNA sequences encoding AID may be used for gene
 CC therapy and the antibodies to the AID protein may be used for diagnosis
 CC and treatment of these disorders. The present sequence represents the
 CC exon 5 genomic DNA sequence of human AID
 CC
 XX
 SQ Sequence 2172 BP; 702 A; 379 C; 465 G; 626 T; 0 U; 0 Other;
 Query Match 77.1%; Score 2172; DB 3; Length 2172;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	623	CCCCGTAATGAGGTGTAATGACTTAACGAGACGCAATTTGTACTTTGGACCTTTGATGACAA	662
Db	1	CCCCGTATGAGGTGTAATGACTTAACGAGACGCAATTTGTACTTTGGACCTTTGATGACAA	60
QY	683	CTTCCAGAAATGTCACACACGATGAATATCTCGCTGGAAGACAGTGGATPAAAAACAGT	742
Db	61	CTTCCAGAAATGTCACACACGATGAATATCTCGCTGGAAGACAGTGGATPAAAAACAGT	120
QY	743	CCCTCAAGTCTTCTGTGTTTATCTTCAACTCTCACTTTCTAGAGTTTACAGAAAA	802
Db	121	CCCTCAAGTCTTCTGTGTTTATCTTCAACTCTCACTTTCTAGAGTTTACAGAAAA	180
QY	803	ATATTTATATACGACTCTTTAAAAAGATCTATGCTTTGAAAATATGAGAGAAACACAGT	862
Db	181	ATATTTATATACGACTCTTTAAAAAGATCTATGCTTTGAAAATATGAGAGAAACACAGT	240
QY	863	CTGGCCAGGGACGCTGCGCAATTGAGTGAAGTTTGAATGCAACATGTCCCTCACTGGGA	922
Db	241	CTGGCCAGGGACGCTGCGCAATTGAGTGAAGTTTGAATGCAACATGTCCCTCACTGGGA	300
QY	923	ATPACAGAACTGCAGACACTGGGAGCACTCTAAAGTGTCAACGTTTCTATGACCTTTA	982
Db	301	ATPACAGAACTGCAGACACTGGGAGCACTCTAAAGTGTCAACGTTTCTATGACCTTTA	360
QY	983	GGTATGATGAGAGCAAGAGTATGATCCTAAAAAGCAATGTGTAGAGGATCAATGTTTTA	1042
Db	361	GGTATGATGAGAGCAAGAGTATGATCCTAAAAAGCAATGTGTAGAGGATCAATGTTTTA	420
QY	1043	TATCAACACTCTTATATATTTGATTCATGTTGAGTTACAGTGGGTATGATATGATTT	1102
Db	421	TATCAACACTCTTATATATTTGATTCATGTTGAGTTACAGTGGGTATGATATGATTT	480
QY	1103	TCATATCTTTTCCCTTGACGTTTACTTTCAGATPACAAACTCTTCATCAGGCCATGA	1162
Db	481	TCATATCTTTTCCCTTGACGTTTACTTTCAGATPACAAACTCTTCATCAGGCCATGA	540
QY	1163	TCATATGAGACTCCTAATGAGATATCTGGGTGATTTGATGCCCAACATCTTCTCCAA	1222
Db	541	TCATATGAGACTCCTAATGAGATATCTGGGTGATTTGATGCCCAACATCTTCTCCAA	600
QY	1223	GCAATTAATCCAAATCATGCGCTGTAATGTTTATATCAGCAGAAACAGTTTATATGTTG	1282
Db	601	GCAATTAATCCAAATCATGCGCTGTAATGTTTATATCAGCAGAAACAGTTTATATGTTG	660
QY	1283	TACAAAAGAAATGTTATGCGTGGGAGTGAAGTATAGACATGATGATCACTTCA	1342
Db	661	TACAAAAGAAATGTTATGCGTGGGAGTGAAGTATAGACATGATGATCACTTCA	720
QY	1343	GCTACTTAATPAAAGATCTTAAATGCGCAGAGACCTGTGAACAAGACCCCTAATA	1402
Db	721	GCTACTTAATPAAAGATCTTAAATGCGCAGAGACCTGTGAACAAGACCCCTAATA	780
QY	1403	TGGGTGATGCTGAATGAGCAATCTCTGGAACCGCAACTCTTTTAAAGAAATGCTCT	1462
Db	781	TGGGTGATGCTGAATGAGCAATCTCTGGAACCGCAACTCTTTTAAAGAAATGCTCT	840
QY	1463	AATTTAGAAACACCAACCACTTACATATCATTAATTAGCAACAAATTTGAGAGAAATG	1522
Db	841	AATTTAGAAACACCAACCACTTACATATCATTAATTAGCAACAAATTTGAGAGAAATG	900
QY	1523	CTTGATGTTGGGAGAGAAAAATCTATGTGCTCTGTGGGTCTTTCATCTCGAAATG	1582
Db	901	CTTGATGTTGGGAGAGAAAAATCTATGTGCTCTGTGGGTCTTTCATCTCGAAATG	960
QY	1583	CCAAACAGGTCAAGTTTGCTACATTTTGTAATGTGTGATGCTTTCGCCAAAGTATAT	1642
Db	961	CCAAACAGGTCAAGTTTGCTACATTTTGTAATGTGTGATGCTTTCGCCAAAGTATAT	1020
QY	1643	TAACTATATPAAAGAGTGTGACAAAAAGATGATPAAAGCTGGGAACCGTGACACGC	1702
Db	1021	TAACTATATPAAAGAGTGTGACAAAAAGATGATPAAAGCTGGGAACCGTGACACGC	1080
QY	1703	TCATGTTCTAGCTCTGGAGAGTGTAGAGGAGATGCTTGAACACAGGTGTTCAA	1762

Dp	1081	TCATAGTTCTAGCTGCTTGGAGGTTAGAGGGGAGATGGCTTGAACAAGGTTCTCA	1140
Qy	1763	GGCCAGCCTGGGCAACATPACAAAGATCCTGTCTCAAAAAAAAAAAAAAAAAA	1822
Dp	1141	GGCCAGCCTGGGCAACATPACAAAGATCCTGTCTCTCAAAAAAAAAAAAAAAAAA	1200
Qy	1823	GAGAGAGGGCCGGGCGTGTGTGTCTACGCGCTGTATCCACACACTTGGAGGCGGAGCC	1882
Dp	1201	GAGAGAGGGCGGGGCGTGTGTGTCTACGCGCTGTATCCACACACTTGGAGGCGGAGCC	1260
Qy	1883	GGGGGANTACACTGTGTGTGACGAGACTTGTGACACGCGCTGGCCAACTGGCAAAACCCGCT	1942
Dp	1261	GGGGGANTACACTGTGTGTGACGAGACTTGTGACACGCGCTGGCCAACTGGCAAAACCCGCT	1320
Qy	1943	CTGTACTCAAAATGCAAAATTTAGCCAGGCGGTGTACAGGCACTGTAAATCCACGCTAC	2002
Dp	1321	CTGTACTCAAAATGCAAAATTTAGCCAGGCGGTGTACAGGCACTGTAAATCCACGCTAC	1380
Qy	2003	TTGGAGGCTGTAGGACAGAGAAATGCTTGAACCCAGAGGTGTGAGTTGCATGATGCTGA	2062
Dp	1381	TTGGAGGCTGTAGGACAGAGAAATGCTTGAACCCAGAGGTGTGAGTTGCATGATGCTGA	1440
Qy	2063	GATGTGCGCCTTGCACCTCCAGCCTGGGCGCAACAAGCAACCTGTGTCTCAGAAAAAAA	2122
Dp	1441	GATGTGCGCCTTGCACCTCCAGCCTGGGCGCAACAAGCAACCTGTGTCTCAGAAAAAAA	1500
Qy	2123	AAAAAAAAAGAGAGAGAGAGAGAGAGACATATTTTGGGAGAGAGATGGGAGACAT	2182
Dp	1501	AAAAAAAAAGAGAGAGAGAGAGAGAGACATATTTTGGGAGAGAGATGGGAGACAT	1560
Qy	2183	TGCAAGAAATGTGCTTATTCACAAATAATGTAAAGAGCCATTAAGGATCCCTATTTTG	2242
Dp	1561	TGCAAGAAATGTGCTTATTCACAAATAATGTAAAGAGCCATTAAGGATCCCTATTTTG	1620
Qy	2243	TCTCTTTGTGTCTATTTGTCCCTAACACTGTCTTGAACGTGAGAAAAATTTTCA	2302
Dp	1621	TCTCTTTGTGTCTATTTGTCCCTAACACTGTCTTGAACGTGAGAAAAATTTTCA	1680
Qy	2303	ATAACCATATCCCTGTGGCGCTTATTAACCTAGCAACCCTGCATGAGATGAGCAGATCC	2362
Dp	1681	ATAACCATATCCCTGTGGCGCTTATTAACCTAGCAACCCTGCATGAGATGAGCAGATCC	1740
Qy	2363	ACAGAAAACTGATGACACACTGTCTTATTTTAACTTATGTATGTACATAATTTGTAA	2422
Dp	1741	ACAGAAAACTGATGACACACTGTCTTATTTTAACTTATGTATGTACATAATTTGTAA	1800
Qy	2423	AGAGTTAAAAATGTGTACTTCACTGATTCATTTATATTTATATTTTGGCGCTATG	2482
Dp	1801	AGAGTTAAAAATGTGTACTTCACTGATTCATTTATATTTATATTTTGGCGCTATG	1860
Qy	2483	ATTTTATTAACAAGATTTCCCTTTTCGATATATTTGAATGGAATGAGGCTCAAAAGCTTCATA	2542
Dp	1861	ATTTTATTAACAAGATTTCCCTTTTCGATATATTTGAATGGAATGAGGCTCAAAAGCTTCATA	1920
Qy	2543	AATTATTAACCTTGAATAATGATTCCTATTAACAACGATATGTAAATTTGAACATTTGACAT	2602
Dp	1921	AATTATTAACCTTGAATAATGATTCCTATTAACAACGATATGTAAATTTGAACATTTGACAT	1980
Qy	2603	GGTGTACGAAGCCATTTCTGTGATTTTAGTAACTTTATGACACGCAATTTGGCTC	2662
Dp	1981	GGTGTACGAAGCCATTTCTGTGATTTTAGTAACTTTATGACACGCAATTTGGCTC	2040
Qy	2663	TGGGTCACTTTCATCAGTTAAATGAATAATGAATAATTTTGAAGCTGTGAAGATTA	2722
Dp	2041	TGGGTCACTTTCATCAGTTAAATGAATAATGAATAATTTTGAAGCTGTGAAGATTA	2100
Qy	2723	TACCAATTAATAATATTAATAAGTATTTATGAGATTTAAATTAATAATCAGTATGAT	2782
Dp	2101	TACCAATTAATAATATTAATAAGTATTTATGAGATTTAAATTAATAATCAGTATGAT	2160
Qy	2783	GGATTAACCTTG 2794	

D 1142 CTCTTCATGAGGCGATGATCTATAGACCTCTCTATAGAGTATCTGGGTATTTGAC 1201
Q 1204 CCCAACCATCTCTCCAAAGCATTAATATGCAATCATGCGCTGATGTTTATCAGCAG 1263
D 1202 CCCAACCATCTCTCCAAAGCATTAATATGCAATCATGCGCTGATGTTTATCAGCAG 1261
Q 1264 AAGATGTTTTTATGTTTATGACAAAAGAAATGTTATGGGTGGGATGAGGTATAGC 1323
D 1262 AAGATGTTTTTATGTTTATGACAAAAGAAATGTTATGGGTGGGATGAGGTATAGC 1321
Q 1324 CATCATGTCACCTTCAAGTACTTAATATAAGATCTTAATAATGGGAGGAGCTGT 1383
D 1322 CATCATGTCACCTTCAAGTACTTAATATAAGATCTTAATAATGGGAGGAGCTGT 1381
Q 1384 GAACAGACACCCCTAATATATGAGTTGATGTCGATAGACAAATCTTGGAAACGCAA 1443
D 1382 GAACAGACACCCCTAATATATGAGTTGATGTCGATAGACAAATCTTGGAAACGCAA 1441
Q 1444 CTCTTTAAGGAGTCCCTAATTTAGAAACACCCAACTTCAATATCATATTAGCA 1503
D 1442 CTCTTTAAGGAGTCCCTAATTTAGAAACACCCAACTTCAATATCATATTAGCA 1501
Q 1504 AACATTTGAAAGAGTGTGTTGATGTTGGGAGAGGAAA 1545
D 1502 AACATTTGAAAGAGTGTGTTGATGTTGGGAGAGGAAA 1543

RESULT 9
AAK81088
ID AAK81088 standard, DNA, 1665 BP.
XX AAK81088;
AC
XX
DT 07-NOV-2001 (first entry)
XX
DS Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35900.
XX
KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX
KN cytosolic; gene therapy; vaccine; metastasis; de.
OS Homo sapiens.
XX
PN WC200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WC-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 18-APR-2000; 2000US-0190076P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216477P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217486P.
PR 14-JUL-2000; 2000US-0218293P.
PR 26-JUL-2000; 2000US-0220930P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232387P.
PR 14-SEP-2000; 2000US-0232388P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236357P.
PR 29-SEP-2000; 2000US-0236358P.
PR 29-SEP-2000; 2000US-0236359P.
PR 29-SEP-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241231P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-024652BP.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249219P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-483426/52.
 PT
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 PS Disclosure; SEQ ID NO 35900; 3072bp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
 CC activity and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK67894 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
 CC represent sequences used in the exemplification of the present invention
 CC
 XX Sequence 1665 BP; 497 A; 343 C; 365 G; 460 T; 0 U; 0 Other;

QY	617	TTTTTGGCCCCGTATGAGTGTGATGACTTACAGAACGATTTTGTACTTTGGACCTTTGA	676
Db	172	CTACGACCCCTGATGAGTGTGATGACTTACAGAACGATTTTGTACTTTGGACCTTTGA	231
QY	677	TAGCAACCTTCAGGAATGTCAACAGCATGAAATATCTGTGCTGAAGACATGTGATAA	736
Db	232	TAGCACTTCAGGAATGTCAACAGCATGAAATATCTGTGCTGAAGACATGTGATAA	291
QY	737	AACAGTCCTCAAGTCTTCTGTGTTTTATTTCTTCAACTCTCACTTCTTAGATTACA	796
Db	292	AACAGTCCTCAAGTCTTCTGTGTTTTATTTCTTCAACTCTCACTTCTTAGATTACA	351
QY	797	GAATAAATATTATATATACACTCTTTAAAAAGATCTATGCTTGAAATATGAGAGAAC	856
Db	352	GAATAAATATTATATACACTCTTTAAAAAGATCTATGCTTGAAATATGAGAGAAC	411
QY	857	AACAGTCGCGCAGAGGACGTGCTCAATTTGGTGTGAGTTTGAATGCAACATTTGCCCTA	916
Db	412	AACAGTCGCGCAGAGGACGTGCTCAATTTGGTGTGAGTTTGAATGCAACATTTGCCCTA	471
QY	917	CTGGGAATTAACAGAACTGCAAGACCTGGGAGCATCTTAAAGTGTCAACGTTTTCTATGA	976
Db	472	CTGGGAATTAACAGAACTGCAAGACCTGGGAGCATCTTAAAGTGTCAACGTTTTCTATGA	531
QY	977	CTTTTAGTAGATGAGAGCAGAAAGTAGATCCTTAAAAAGCATGTTGAGAGATCAATG	1036
Db	532	CTTTTAGTAGATGAGAGCAGAAAGTAGATCCTTAAAAAGCATGTTGAGAGATCAATG	591
QY	1037	TTTTTATATCAACATCCTTTATTTATTTGATTCATTTGAGTTAAAGTGTGTATGTATA	1096
Db	592	TTTTTATATCAACATCCTTTATTTATTTGATTCATTTGAGTTAAAGTGTGTATGTATA	651
QY	1097	GATTTTTCATCTCTTTCCCTGAGCTTTACTTTCAAGTAAACAACATCTCTTCATCAGG	1156
Db	652	GATTTTTCATCTCTTTCCCTGAGCTTTACTTTCAAGTAAACAACATCTCTTCATCAGG	711
QY	1157	CCATGATCTATAGGACCTCTCTAATGAGAGTATCTGGGTGATTTGACCCCAACATCTC	1216
Db	712	CCATGATCTATAGGACCTCTCTAATGAGAGTATCTGGGTGATTTGACCCCAACATCTC	771
QY	1217	TCCAAGCATTAATATCCAAATCATGCGCTATGTTTTAATCAGACAGACATGTTTTTA	1276
Db	772	TCCAAGCATTAATATCCAAATCATGCGCTATGTTTTAATCAGACAGACATGTTTTTA	831
QY	1277	TGTTTGTACAAABAAGATGTTATGTGGTGGGATGAGGATAGACATGCAATGCTAC	1336
Db	832	TGTTTGTACAAABAAGATGTTATGTGGTGGGATGAGGATAGACATGCAATGCTAC	891
QY	1337	CTTCAGGCTACTTTAATTAAGAGATCTTAAAAATGGGACGAGAGACTGTGAACAACACC	1396
Db	892	CTTCAGGCTACTTTAATTAAGAGATCTTAAAAATGGGACGAGAGACTGTGAACAACACC	951
QY	1397	TAAATATGGGTGATGTCTGATAGCAAAATCTTCTGAAAACGCAAACTCTTTAAGGAA	1456
Db	952	TAAATATGGGTGATGTCTGATAGCAAAATCTTCTGAAAACGCAAACTCTTTAAGGAA	1011
QY	1457	GTCCTTAATTTAGAAAACCCCAAACTTCACATATCATTAATTAGCAACAAATTGGAAG	1516
Db	1012	GTCCTTAATTTAGAAAACCCCAAACTTCACATATCATTAATTAGCAACAAATTGGAAG	1071
QY	1517	AAGTTGCTGTGAATGTGGGGAGAGAAAATCTATTTGGCTCTCGGGGCTCTTATCTCA	1576
Db	1072	AAGTTGCTGTGAATGTGGGGAGAGAAAATCTATTTGGCTCTCGGGGCTCTTATCTCA	1131
QY	1577	GAATATGCCAATCAGGTCAGGTTTGTACATTTTGTATGTGTGATGCTTCTCCCAAG	1636
Db	1132	GAATATGCCAATCAGGTCAGGTTTGTACATTTTGTATGTGTGATGCTTCTCCCAAG	1191
QY	1637	GTATATTTACTATATTAAGAGATTGTGACAAACAGATGATTAAGCTCGCAACCTGTGC	1696
Db	1192	GTATATTTACTATATTAAGAGATTGTGACAAACAGATGATTAAGCTCGCAACCTGTGC	1251

PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259578P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 P1 Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Disclosure; SEQ ID NO 35901; 3071bp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SO Sequence 574 BP; 201 A; 80 C; 85 G; 208 T; 0 U; 0 Other;

Query Match 20.4%; Score 574; DB 4; Length 574;
 Best Local Similarity 100.0%; Pred. No. 1.5e-90;
 Matches 574; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2211 AATGTAAGAGCCATAAGGATCCCTATTGTCCTTTGGTGTCTATTGTCCTTAAC 2270
 DB 1 AATGTAAGAGCCATAAGGATCCCTATTGTCCTTTGGTGTCTATTGTCCTTAAC 60
 QY 2271 AACTGCTTTGACGTGAGAAAAATATTGAGATTAACCATATCCCTGCGCTATTAC 2330
 DB 61 AACTGCTTTGACGTGAGAAAAATATTGAGATTAACCATATCCCTGCGCTATTAC 120
 QY 2331 TAGCAACCCCTGCATGAAGATGAGAGATCCACAGAAACTTGATGCAACGTCT 2390
 DB 121 TAGCAACCCCTGCATGAAGATGAGAGATCCACAGAAACTTGATGCAACGTCT 180
 QY 2391 TATTTAATCTTATGATCAATAGTTGTAAAGAGTTAAAAATGTTACTGATGAT 2450
 DB 181 TATTTAATCTTATGATCAATAGTTGTAAAGAGTTAAAAATGTTACTGATGAT 240
 QY 2451 CATTTAATTTTATTTATTTTGGTCTAATGATTTTATTAACATGATTTCTTTTCT 2510
 DB 241 CATTTAATTTTATTTATTTTGGTCTAATGATTTTATTAACATGATTTCTTTTCT 300
 QY 2511 GATATATGGAATGAGAGCTCAAGCTTCAATTAATTAATTAATTAATTAATTAAT 2570
 DB 301 GATATATGGAATGAGAGCTTCAAGCTTCAATTAATTAATTAATTAATTAATTAAT 360
 QY 2571 AACACGTATGTAATTTAAATTCATTCAGTATGAGTGTACGAAGCCATTTCTTGATT 2630

DB 361 AACACGTATGTAATTTAAATTCATTCAGTATGAGTGTACGAAGCCATTTCTTGATT 420
 QY 2631 TTATGAACTTTTATGACAGCAAAATTTGCTTGCTCACTTTCAATCAATTAAAT 2690
 DB 421 TTATGAACTTTTATGACAGCAAAATTTGCTTGCTCACTTTCAATCAATTAAAT 480
 QY 2691 GATATATGGAATGAGAGCTGGAAGTAAATTCGAAATTAATTAATTAATTAATTAAT 2750
 DB 481 GATATATGGAATGAGAGCTGGAAGTAAATTCGAAATTAATTAATTAATTAATTAAT 540
 QY 2751 TATATGAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2784
 DB 541 TATATGAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 574
 RESULT 13
 AAK61819
 ID AAK61819 standard; cDNA; 429 BP.
 XX
 AC AAK61819;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:6879.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 XX cytostatic; gene therapy; vaccine; metastasis; ss.
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001354.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186359P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 11-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225457P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226641P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235444P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241211P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241866P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251038P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254037P.
PR 05-JAN-2001; 2001US-02595678P.

(HOMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-483426/52.
P-PSDB; AAM89038.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and metastasis.

PS Claim 1; SEQ ID NO 6879; 3071bp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK67694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 429 BP; 144 A; 79 C; 97 G; 109 T; 0 U; 0 Other;

Query March 15.2%; Score 429; DB 4; Length 429;
Best Local Similarity 100.0%; Pred. No. 2.2e-65;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1386 ACAAGACCCCTATATATGAGTGTCTGAGTAGCAATCTTGTGAAAGCAACT 1445
DB 1 ACAAGACCCCTATATATGAGTGTCTGAGTAGCAATCTTGTGAAAGCAACT 60
QY 1446 CTTTAAAGAGAGTCCCTAATTAGAAACACCAACCACTTACATATCATTAATAGCAA 1505
DB 61 CTTTAAAGAGAGTCCCTAATTAGAAACACCAACCACTTACATATCATTAATAGCAA 120
QY 1506 CATTGGAGAGAGTCTCTGATGTGGGAGAGAGAAATCTATTGGCTCTGTGGGTC 1565

Db 121 CAATTGGAAGAAATGCTTGAATGTTGGGAGAGAAAACTATTGGCTCTGCGGCTC 180
QY 1566 TCTTCATCTCAGAAATGCCATCAGTCAAGGTTTGCTACATTTGTATGTGTGATGC 1625
Db 181 TCTTCATCTCAGAAATGCCATCAGTCAAGGTTTGCTACATTTGTATGTGTGATGC 240
QY 1626 TTTCTCCCAAGGTATTTACTATATTAAGAGAGTTTGACAAAACAGATGATAAGCTG 1685
Db 241 TTTCTCCCAAGGTATTTACTATATTAAGAGAGTTTGACAAAACAGATGATAAGCTG 300
QY 1686 CGAACCGTGGACACGCTCATAGTCTAGCTGCTGGAGAGTTGAGAGGAGATGCGCT 1745
Db 301 CGAACCGTGGACACGCTCATAGTCTAGCTGCTGGAGAGTTGAGAGGAGATGCGCT 360
QY 1746 TGAACACAGGTGTTCAAGGCGAGCTG3GCAACATAACAAGATCTGTCTCAAAAAA 1805
Db 361 TGAACACAGGTGTTCAAGGCGAGCTG3GCAACATAACAAGATCTGTCTCAAAAAA 420
QY 1806 AAAAAAAAA 1814
Db 421 AAAAAAAAA 429
RESULT 14
AAL02789 standard; DNA; 13862 BP.
XX AAL02789;
AC AAL02789;
XX 21-NOV-2001 (first entry)
DT 21-NOV-2001 (first entry)
XX Human reproductive system related antigen DNA SEQ ID NO: 5477.
DE Human reproductive system related antigen; reproductive system disorder;
XX Human; reproductive system related antigen; reproductive system disorder;
KM cancer; gene therapy; ds.
XX Homo sapiens.
OS Homo sapiens.
XX WO200155320-A2.
PV WO200155320-A2.
XX 02-AUG-2001.
PD 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001339.
PF 17-JAN-2001; 2001WO-US001339.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0230963P.
PR 26-JUL-2000; 2000US-0230964P.
PR 14-AUG-2000; 2000US-0234518P.
PR 14-AUG-2000; 2000US-0234519P.
PR 14-AUG-2000; 2000US-0235213P.
PR 14-AUG-2000; 2000US-0235214P.
PR 14-AUG-2000; 2000US-0235266P.
PR 14-AUG-2000; 2000US-0235267P.
PR 14-AUG-2000; 2000US-0235268P.
PR 14-AUG-2000; 2000US-0235270P.
PR 14-AUG-2000; 2000US-0235447P.
PR 14-AUG-2000; 2000US-0235757P.
PR 14-AUG-2000; 2000US-0235758P.

PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227183P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229515P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231245P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236328P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0236803P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249298P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-46570/50.
 XX
 PT Isolated nucleic acid molecule encoding a reproductive system antigen is
 used in preventing, treating or ameliorating a medical condition.
 XX
 XX
 XX Disclosure; SEQ ID NO 5477; 1297bp + Sequence Listing; English.
 CC
 CC The present invention provides the protein and coding sequences of a
 number of human reproductive system related antigens. These can be used
 in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a genomic sequence encoding a
 protein of the invention
 CC
 XX
 XX Sequence 13862 BP; 3188 A; 3560 C; 4465 G; 2649 T; 0 U; 0 Other;

Query Match 10.5%; Score 295.2; DB 4; Length 13862;
 Best Local Similarity 74.6%; Pred. No. 5.9e-42;
 Matches 397; Conservative 0; Mismatches 133; Indels 2; Gaps 2;

QY 1667 AAACAGATATTAAGCTGGCAACGCTGACACGCTCATAGTTCTAGCTGCTGGAGG 1726
 Db 1508 AAACCTTAATAGCTGGCTGGCTGCTGCTCATGCTCATGCTGATCCAGACATTGGAGG 1567
 QY 1727 TTGAGGAGGAGGAGTGGCTTGAACACAGCTTCAAGGCGCAGCTGGGACATATACAG 1786
 Db 1568 CTGAGGAGGAGGAGTCACTTGAAGTCAAGGATTGAGACCAAGCTGACCACTGGCGAA 1627
 QY 1787 ATCTGTCTCTCAAAAAAAAAAAAAAAAAAGAAAGAGAGGCGCGGCGCTGGTGGC 1845
 Db 1628 ACCGCATCTCTACATTAATAATCAAAAGTTAGCCAGGCAATGGCGGCGATGGTGGC 1687
 QY 1846 TCACGCTGTAAATCCAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGGTCAGA 1905
 Db 1688 TCATGCTGTAAATCCAGCACTTTGGAGGCGGAGGCGGATCACTGTGGTCAGA 1747

QY 1906 GTTTGAGACGACCTGGCGCAATGGCAAAACCCGCTGTGACTCAAAATGCAAAATTA 1965
 Db 1748 GTTCAAGACGACCTGGCAAAATGGCAAAACCCGCTGTGACTCAAAATTA 1807
 QY 1966 GCCAGGCGTGTAGCAGGACCTGTAATCCAGCTACTTGGAGGCTGAGGCGAGAAAT 2025
 Db 1808 GCCAGGCTTATGTAGGCGCTATTAATCCAGCTACTTGGAGGCTGAGGCGAGAAAT 1867
 QY 2026 CGCTTGAACCCAGAGAGTGGAGGTTGACATGATGATCGCTGCTGACTCCAGCC 2085
 Db 1868 CTTTGAAGCCAGAGGCGAGGTTTGACCGAGCGAGATTGTGCCACTGCATCCAGCC 1927
 QY 2086 TGGGCGACAAAGCAAGACTTGTCTCAGAAAAAAGAGAGAGAGAGAGAA 2145
 Db 1928 TGGGCGATTA-AGCAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAA 1986
 QY 2146 AGAGACAAATTTTGGGAGAGAGATGGGAGACATTGCAAGGAAATTGG 2197
 Db 1987 AGAAAAATTAGCCAGCGTGTGATGACCTGTAGTCTTAGCTACTTGG 2038

RESULT 15
 AA07516
 ID AA07516 standard; DNA; 13862 BP.

XX AA07516;

XX 21-NOV-2001 (first entry)

DE Human reproductive system related antigen DNA SEQ ID NO: 10204.

KW Human; reproductive system related antigen; reproductive system disorder;
 cancer; gene therapy; ds.

XX Homo sapiens.

XX WO20015320-A2.

XX 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US001339.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217486P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

```
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228928P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 05-SEP-2000; 2000US-0230437P.
PR 05-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231422P.
PR 08-SEP-2000; 2000US-0231433P.
PR 08-SEP-2000; 2000US-0231443P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 25-SEP-2000; 2000US-0235844P.
PR 25-SEP-2000; 2000US-0235845P.
PR 27-SEP-2000; 2000US-0235836P.
PR 27-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241211P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465570/50.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
PT used in preventing, treating or ameliorating a medical condition.
XX
XX Disclosure; SEQ ID NO 10204; 1297pp + Sequence listing; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention
XX
XX
SQ Sequence 13862 BP; 3188 A; 3560 C; 4465 G; 2649 T; 0 U; 0 Other;

Query Match 10.5%; Score 295.2; DB 4; Length 13862;
Best Local Similarity 74.6%; Pred. No. 5.9e-42;
Matches 397; Conservative 0; Mismatches 133; Indels 2; Gaps 2;

QY 1667 AAACGAATGATAAGCTGGGAACCGTGGACGACGCGCTGATGTTAGTGGTGGAGG 1726
DB 1508 AACTTATATAGCTGGGTGGGCTGTGTGCTCATGCTGTGATCCAGCATTTGGAGG 1567
QY 1727 TTGAGAGAGAGATGCTTGAACACAGGTGTTCAAGGCCAGCTGGGCAATTAACAG 1786
DB 1568 CTGAGCAGAGAGATCACTTGAAGTCAAGGTTGAGACCAAGCTGGCCAAATGGCGAA 1627
QY 1787 ATCTGTCTCTCAAAAAAAAAAAAAAAAAAAGAAAG- GAGAGGCGCGGCGCTGTGGC 1845
DB 1628 ACCCATCTCTACTAAAAAATACAAAGTTAGCCAGGATGGCGCGCATGTGGGTGGC 1687
QY 1846 TCAGCGCTGTAAATCCAGCATTTGGAGGCGGAGCGGCGGATACCTGATGGTGAAGA 1905
DB 1688 TCATGCTGTATATCCAGCATTTGGAGGCGGAGCGGCGGATACCTGATGGTGAAGA 1747
```

```

QY 1906 GTTTGAGACCAAGCTGGCCACATGSCAAACCCTGCTGTAATAAATGAAAATTA 1965
Db 1748 GTTCAGACCAAGCTGGACACATGSCAAACCCTGCTGTAATAAATGAAAATTA 1807
QY 1966 GCCAGGCGTGTAGCAGGACCTGTATCCAGCTACTTGGAGGCTGAGGACGAGAAAT 2025
Db 1808 GCCAGGCTTATGTAGTGGCGCCTATATATCCAGCTACTTGGAGGCTGAGGACGAGAAAT 1867
QY 2026 CGCTTGAAACCCAGAGAGTGGAGGTTGCAAGTAAGCTGAATCGTGCCTGTCACCTCAGCC 2085
Db 1868 CGTTGAAAGCCAGAGGCGAGATTGTCACCGAGCGAGATGTGCCACTGCACCTCAGCC 1927
QY 2086 TGGGCGACAGAGCAAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGAA 2145
Db 1928 TGGGCGATTA-AGCGAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGAA 1986
QY 2146 AGAGAACATATTTGGAGAGAAAGATGGGGAAGCATTCAGAGAAATTGTG 2197
Db 1987 AGAAAAAATTAGCGAGCGTGTGGCATGCACTGTAGTCTTACTTGGG 2038

```

Search completed: March 7, 2004, 06:55:59
 Job time : 1065 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 06:31:31 ; Search time 217 Seconds

(without alignments)
7206.689 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818
Sequence: 1 agagaacatcatatga.....aaaaaaaaaaaaaaaa 2818

Scoring table: IDENTITY NUC
Gapco 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	280	9.9	15297	4 US-09-817-180-3	Sequence 3, Appli
2	280	9.9	15297	4 US-10-003-295-3	Sequence 3, Appli
3	274.2	9.7	11811	3 US-09-078-294-7	Sequence 7, Appli
4	269.8	9.6	43950	4 US-09-735-934A-3	Sequence 3, Appli
5	269.8	9.6	43950	4 US-10-060-332-3	Sequence 3, Appli
6	269.2	9.6	99500	4 US-09-798-096-10	Sequence 10, Appli
7	267.4	9.5	59065	4 US-09-813-817-3	Sequence 3, Appli
8	267.4	9.5	59065	4 US-09-978-197-3	Sequence 3, Appli
9	265.6	9.4	9365	4 US-09-608-285A-8	Sequence 8, Appli
10	265.6	9.4	9365	4 US-09-350-836B-8	Sequence 8, Appli
11	265.6	9.4	9365	4 US-09-370-265-8	Sequence 8, Appli
12	265.6	9.4	9365	4 US-09-557-800C-8	Sequence 8, Appli
13	265.6	9.4	9365	4 US-09-570-625A-8	Sequence 8, Appli
14	265.6	9.4	14747	4 US-09-608-285A-42	Sequence 42, Appli
15	265.6	9.4	14747	4 US-09-557-800C-42	Sequence 42, Appli
16	265.6	9.4	15977	4 US-09-608-285A-59	Sequence 59, Appli
17	262.2	9.3	21234	4 US-09-810-671-3	Sequence 3, Appli
18	262.2	9.3	21234	4 US-10-109-854-3	Sequence 3, Appli
19	260.4	9.2	63000	4 US-09-780-172-18	Sequence 18, Appli
20	259.2	9.2	3885	4 US-08-688-145-1	Sequence 3, Appli
21	258.6	9.2	43950	4 US-09-735-934A-3	Sequence 3, Appli
22	258.6	9.2	43950	4 US-10-060-332-3	Sequence 3, Appli
23	257.6	9.1	99500	4 US-09-798-096-10	Sequence 10, Appli
24	257.2	9.1	10684	4 US-08-618-100B-3	Sequence 3, Appli
25	256.8	9.1	8174	1 US-07-914-281-5	Sequence 5, Appli
26	256.8	9.1	8174	1 US-08-393-246-5	Sequence 5, Appli
27	256.8	9.1	8174	1 US-08-525-058A-5	Sequence 5, Appli

C 28	256.8	9.1	8174	2 US-08-696-731-5	Sequence 5, Appli
C 29	256.8	9.1	8174	5 US-09-042-531-5	Sequence 5, Appli
C 30	256.8	9.1	8174	5 PCT-US91-00895-3	Sequence 3, Appli
C 31	256.8	9.1	29629	4 US-09-729-995-3	Sequence 3, Appli
C 32	256.8	9.1	29629	4 US-10-135-689-3	Sequence 3, Appli
C 33	254.8	9.0	118067	4 US-09-497-855A-32	Sequence 32, Appli
C 34	254.4	9.0	98844	4 US-09-791-211-10	Sequence 10, Appli
C 35	253.8	9.0	16053	4 US-09-801-052-3	Sequence 3, Appli
C 36	253.8	9.0	16053	4 US-10-020-121-3	Sequence 3, Appli
C 37	253.2	9.0	14581	4 US-08-520-373D-4	Sequence 4, Appli
C 38	253.2	9.0	22481	4 US-08-367-841A-43	Sequence 43, Appli
C 39	253.2	9.0	22481	5 PCT-US95-07201-43	Sequence 43, Appli
C 40	253.2	9.0	22484	4 US-09-875-223-2	Sequence 2, Appli
C 41	253.2	9.0	22484	4 US-09-875-114-2	Sequence 2, Appli
C 42	252.2	8.9	7210	2 US-08-257-963B-10	Sequence 10, Appli
C 43	252.2	8.9	7210	4 US-08-367-841A-10	Sequence 10, Appli
C 44	252.2	8.9	7210	5 PCT-US95-07201-10	Sequence 10, Appli
C 45	252.2	8.9	14581	4 US-08-520-373D-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-817-180-3
Sequence 3, Application US/09817180
Patent No. 6340584
GENERAL INFORMATION:
APPLICANT: GAN, Weiniu et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THREEOF
FILE REFERENCE: C1001183
CURRENT APPLICATION NUMBER: US/09/817,180
CURRENT FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 15297
TYPE: DNA
ORGANISM: Human
US-09-817-180-3

Query Match 9.9%; Score 280; DB 4; Length 15297;
Best Local Similarity 74.9%; Pred. No. 3e-52;
Matches 370; Conservative 0; Mismatches 110; Indels 14; Gaps 1;

QY	1641	ATTAACTATATTAAGAGATTGACAAACAGATATTAAGCTGGACCGTGGACAC	1700
DB	11651	ACTCCCTTAATGCGCACTCCACACAGGAGATTAAGATTAACCTGGCAGTGTCCAC	11710
QY	1701	GCTCATAGTCTGCTGCTGCTGGAGGTGAGAGGAGAGATGCTTGAACACAGGTTC	1760
DB	11711	GCTGTATATCCAGCACTTTGGAGGCTGAGCTGATCACTTGAGCCAGAGTTC	11770
QY	1761	AAGGCGAGCTGGGCAACATPAACAAGTCCTGCTCTCAAAAAAAAAAAAAAAAAA	1820
DB	11771	AAGATCAGCTTGGACACAGAGTGAATCTCATCAAAAAAAAAAAAAAAAAA	11825
QY	1821	AAGAGAGAGGCGCGGGGCTGAGCTCAGCCTGTATCCAGCACTTTGGAGGCGGAG	1880
DB	11826	-----GACTGGGACGCTGGCTCAGCCTGTATCCAGCACTTTGGAGGCGGAG	11876
QY	1881	CCGAGCGGATCAGCTGTGTGTGAGAGTTGAGCCAGCTGGGCAACATGGCAAAACCC	1940
DB	11877	GCAAGTGAATCAGCTGTGTGTGAGAGTTGAGCCAGCAACATGGTAAACCCC	11936
QY	1941	GTCGTACTCAAAATGCAAAATTTAGCAGGCGTGGAGGACACCTGTATCCAGCT	2000
DB	11937	ATCTCTTAATAAATTAATTAATTTAGCAGGCAATGGTGGACCTGTATCCAGCT	11996
QY	2001	ACTTGGAGGCTGAGGAGGAGATGCTTGAACCAAGAGGTGAGGTTGACATAGCT	2060

Db 11997 ACTTGAGAGCTGAGGTGGAGAAATTGCTTGAACCCAGAGCGAGGCTGACAGTGAACC 12056
QY 2061 GAGATCGTCCGCTTGCTGCTCCAGCTCGGCGCAAGAGCAAGACTCTGCTCAGAAAAA 2120
Db 12057 GAGATTGTCCACTGCACTCCAGCTCGGCGCAAGAGTGAATCTCATCTCAAAAAA 12116
QY 2121 AAAAAAAAAAGAGA 2134
Db 12117 CCAAAAAACAAAAA 12130

RESULT 2

US-10-003-295-3
; Sequence 3, Application US/10003295
; Patent No. 6686187
; GENERAL INFORMATION:
; APPLICANT: GAT, Weidm et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-3

Query Match 9.9%; Score 280; DB 4; Length 15297;
Best Local Similarity 74.9%; Pred. No. 3e-52;
Matches 370; Conservative 0; Mismatches 110; Indels 14; Gaps 1;

QY 1641 ATTAATCTATTAAGAAGTTGTGACAAAAACAAATGATTAAGCTGGAAACCTGGCACAC 1700
Db 11651 ACTCCCTTAATGCAACCTTCCACAGAGAGATTAAGATTAACCTGGCCAGTTGCTCAC 11710
QY 1701 GCTCATAGTTCTAGCTGCTTGGAGGTTGAGAGGAGGATGGCTTGAACACAGGTTTC 1760
Db 11711 GCTGTGATCTCCAGCACTTTGGAGGCTGAGCTGGTGAATCACTTGAACCCAGAGTTTC 11770
QY 1761 AAGCCAGCTCGGCAACATTAACAAGATCTGTCTCTCAAAAAAATAAAAAAAGA 1820
Db 11771 AAGATCAGCTTGAACAACAGTGAATCTCTTACAAAAAATAAAAAAATA----- 11825
QY 1821 AAGAGAGAGGCGCGGCGTGTGCTGACGCTGTATCCAGCACTTTGGAGGCGGAG 1880
Db 11826 -----GACTGGGACCGGTGGTCACTCTGTATCCAGCACTTTGGAGGCGGAG 11876
QY 1881 CCGGCGGATCACCTGTGTGACAGAGTTGAGACCGAGCTGGCCACATGGCAAAACCC 1940
Db 11877 GCAAGTGTGATCAGCTGTGTGAGAGTTGAGACCGAGCAACCAATGTGAAACCC 11936
QY 1941 GTCTGTACTCAAAATGCAAAATTAAGCAGGCGTGTGAGAGCACTGTAAATCCAGCT 2000
Db 11937 ATCTCTACTAAATTAACAAATTAAGCAGGCAATGTGTGACAGCTGTAAATCCAGCT 11996
QY 2001 ACTTGGAGGCTGAGGAGAGAGATCGTTGAACCAAGAGGTGAGGTTGCAATAGCT 2060
Db 11997 ACTTGGAGGCTGAGGAGAGATCGTTGAACCAAGAGGCGGAGGCTGCAATAGGCT 12056
QY 2061 GAGATCGTCCGCTTGCTGCTCCAGCTCGGCGCAAGAGCAAGCTGTCTCAGAAAAA 2120
Db 12057 GAGATTGTCCACTGCACTCCAGCTCGGCGCAAGAGTGAATCTCATCTCAAAAAA 12116
QY 2121 AAAAAAAAAAGAG 2134
Db 12117 CCAAAAAACAAAAA 12130

RESULT 3

US-09-078-294-7/c
; Sequence 7, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 11811
; TYPE: DNA
; ORGANISM: BAC-F2 contig 3
US-09-078-294-7

Query Match 9.7%; Score 274.2; DB 3; Length 11811;
Best Local Similarity 79.7%; Pred. No. 5.2e-51;
Matches 350; Conservative 0; Mismatches 83; Indels 6; Gaps 2;

QY 1692 GTGGCAGGCTCATAGTTGTTAGCTGCTGGGAGGTTGAGAGGAGAGATGGCTTGAACA 1751
Db 6663 GTGGCTATGCTGTGTAATCCAGCACTTTGGAGGCGCGAGGTGGAGAAATGACTTGAAGT 6604
QY 1752 CAGGTTTCAAGGCCAGCTCGGCAACATTAACAAGATCTGTCTCAAAAAAATAAAAAA 1811
Db 6603 CAGAGTTCAAGACAGCTCGGCAACAGTGAAGCTCATCTCAAAAAATACATTTA 6544
QY 1812 AAAAAAAAAAGAGAGAGGCGCGGCGTGTGCTCAAGCTGTATCCAGCACTTTGG 1871
Db 6543 AAGATTAG-----CTGGGCGCAGGTGTGTGCGCAGCCTGTATCCAGCCTTTGG 6489
QY 1872 GAGGCCAGCGCGGAGATCACTGTGTGAGAGTTGAGACCACTGGCCAAACATGG 1931
Db 6488 GAGGTCAAGGTGTGTATCACTGAGGTGAGAGTTGAGGCCAGCTGGCCAAACATGG 6429
QY 1932 CAAGACCCCTGTGTAATCAAAATGC-AAAATTAAGCAGGCGTGTGAGCAAGCTGT 1990
Db 6428 TGAAACCCCTGTCTACTTAATAAATAAATAAATTAAGCAGGCGTGTGCGGCGCTGT 6369
QY 1991 AATCCAGCTACTTGGAGGCTGAGGAGAGATCGTTGAACCAAGAGGTGAGGTT 2050
Db 6368 AATCCAGCTACTGAGGAGGCTGAGGAGAGAAATCACTTGAACCAAGAGAGGAGTT 6309
QY 2051 GCAATAGCTGAGATCGTCCGCTTGCTGCACTCCAGCTGGGCGCAAGAGCAAGCTGTTC 2110
Db 6308 GCACTGAGCCAAAGTCAATGCAATGCACTCCAGCTGGGCGCAAGAGCAAACTCTCA 6249
QY 2111 TCAGAAAAAATAAAAAA 2129
Db 6248 AAAAAAAAAAAAAA 6230

RESULT 4

US-09-735-934A-3/c
; Sequence 3, Application US/09735934A
; Patent No. 6372468
; GENERAL INFORMATION:
; APPLICANT: Li, Jia-Yin et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000851
; CURRENT APPLICATION NUMBER: US/09/735,934A
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 43950
; TYPE: DNA

Db 43518 TCTGTATCCAGCTACTGCGGAGGCTGAGGAGGAAATCGCTTGAACCTGGAGGCTGG 43459
QY 2046 AGGTTSCAGTAAGCTAGATCGTGGCGTTGCACTCCAGCTGGCGAGCAAGAGACT 2105
Db 43458 AGGTTSCAGTAAGCTAGATCGTGGCGTTGCACTCCAGCTGGCGAGCAAGAGACT 43399
QY 2106 CTGTCTCAGAAAAAAGAAAAAAGAGAGAGAGAGAGAAAG 2149
Db 43398 TTGTCTCAGAAAAAAGAAAAAAGAGAGAGAGAGAGAAAGG 43355

RESULT 7

US-09-813-817-3/c
; Sequence 3, Application US/09813817
; Patent No. 6340583
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEIN, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001178
; CURRENT APPLICATION NUMBER: US/09/813,817
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-813-817-3

Query Match 9.5%; Score 267.4; DB 4; Length 59065;
Best Local Similarity 78.2%; Pred. No. 3.1e-49;
Matches 351; Conservative 0; Mismatches 86; Indels 12; Gaps 2;

QY 1680 AAGCTGGAACCGTGGACACGCTCAATGTTCTAGCTGCTGGAGGTTGAGAGGAGG 1739
Db 23369 AGCCAGGCAATAGTGGCTCAATGCTATTAATCCAGC-ATTGAGAGGCGCAAGAGGAGG 23311
QY 1740 ATGGCTTGAACACAGGTGTTCAAGGCGCAGCTGGGCAATACAAAGATCTGTCTCA 1799
Db 23310 ACTGCTTGAAGCCAGAAAGTTTGAGACACGCTGGGTAACAAACCAAGACCTTATCTTAC 23251
QY 1800 AAAAAAAGAAAAAAGAAAAAGAGAGAGGCGCGGCGTGGTCAAGCTGTAAATC 1859
Db 23350 TAAAAAATAGTCAATCA-----GGCCAGTCAATGTGGCTCAGCCTGTATC 23202
QY 1860 CAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGATCAGAGTTTGAAGCAGC 1919
Db 23301 CAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGATCAGAGTTTGAAGTTGAGCAGC 23142
QY 1920 TGGCCAAATGAGCAAAACCCGCTGTCTGATCTCAAAATGCAAAATTTGCGAGCGCTGTAG 1979
Db 23141 TGACCAACATGAGAAACCTGTCTCTACTTAAATAATTAATAATTTGCGAGCGCTGTAG 23082
QY 1980 CAGGCACTGTATCCAGCTACTTGGAGGCTGAGAGGAGAAATCGCTTGAACCCAGG 2039
Db 23081 CAGTGTCTGTATCCAGCTACTTGAAGGCTGAGAGGAGAAATCGCTTGAACCCAGG 23022
QY 2040 AGGTGAGGTTGAGTGAAGTGAAGTGTGCGCTTGCACCTCCAGCTGGGCGAGCAAGAGC 2099
Db 23021 AGGAGAGGTTGGCGGTGAGTGTGCTGTGCTTCCATTCAGCCTGGGCAAGAGAGT 22962
QY 2100 AAGACTGTGTCTCAGAAAAAAGAAAAA 2128
Db 22961 GAAACTGTGTCTCAGAAAAAAGAAAAA 22933

RESULT 8
US-09-978-197-3/c
; Sequence 3, Application US/09978197
; Patent No. 6403353
; GENERAL INFORMATION:

; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEIN, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001178D10
; CURRENT APPLICATION NUMBER: US/09/978,197
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/813,817
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-978-197-3

Query Match 9.5%; Score 267.4; DB 4; Length 59065;
Best Local Similarity 78.2%; Pred. No. 3.1e-49;
Matches 351; Conservative 0; Mismatches 86; Indels 12; Gaps 2;

QY 1680 AAGCTGGAACCGTGGACACGCTCAATGTTCTAGCTGCTGGAGGTTGAGAGGAGG 1739
Db 23369 AGCCAGGCAATAGTGGCTCAATGCTATTAATCCAGC-ATTGAGAGGCGCAAGAGGAGG 23311
QY 1740 ATGGCTTGAACACAGGTGTTCAAGGCGCAGCTGGGCAATACAAAGATCTGTCTCA 1799
Db 23310 ACTGCTTGAAGCCAGAAAGTTTGAGACACGCTGGGTAACAAACCAAGACCTTATCTTAC 23251
QY 1800 AAAAAAAGAAAAAAGAAAAAGAGAGGCGCGGCGTGGTCAAGCTGTAAATC 1859
Db 23350 TAAAAAATAGTCAATCA-----GGCCAGTCAATGTGGCTCAGCCTGTAAATC 23202
QY 1860 CAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGATCAGAGTTTGAAGCAGC 1919
Db 23301 CAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGATCAGAGTTTGAAGTTGAGCAGC 23142
QY 1920 TGGCCAAATGAGCAAAACCCGCTGTCTGATCTCAAAATGCAAAATTTGCGAGCGCTGTAG 1979
Db 23141 TGACCAACATGAGAAACCTGTCTCTACTTAAATAATTAATAATTTGCGAGCGCTGTAG 23082
QY 1800 CAGGCACTGTATCCAGCTACTTGGAGGCTGAGAGGAGAAATCGCTTGAACCCAGG 2039
Db 23081 CAGTGTCTGTATCCAGCTACTTGAAGGCTGAGAGGAGAAATCGCTTGAACCCAGG 23022
QY 2040 AGGTGAGGTTGAGTGAAGTGAAGTGTGCGCTTGCACCTCCAGCTGGGCGAGCAAGAGC 2099
Db 23021 AGGAGAGGTTGGCGGTGAGTGTGCTGTGCTTCCATTCAGCCTGGGCAAGAGAGT 22962
QY 2100 AAGACTGTGTCTCAGAAAAAAGAAAAA 2128
Db 22961 GAAACTGTGTCTCAGAAAAAAGAAAAA 22933

RESULT 9

US-09-608-285A-8/c
; Sequence 8, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11


```

; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc_feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-557-800C-8

Query Match          9.4%; Score 265.6; DB 4; Length 9365;
Best Local Similarity 79.0%; Pred. No. 3.8e-49;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1765 CCAGCCTGGGCAACATACAGATCCTGTCTCTCAAAAAAAAAAAAAAAAAAGAAAG-1823
DB 7802 CCAGCATGGGTGACAGAGCAAGACCTGTCTCAAAAAAAAAAAAAAAAAAGAAAG 7743

QY 1824 -----AGAGAGGCGCGGCGTGTGCTCAAGCTGTATCCAGACCTTGGAGG 1875
DB 7742 AAMGAMCAAAAAAAAAAGGTGAGGTGTGCTCACTCTTAATCTTGAAGCTTTGGAGG 7683

QY 1876 CCGAGCCGCGGCGGATCACTGTGTGTCAGAGTTGAGACCAAGCCTGCGCAACATGCGAAA 1935
DB 7682 CYGAGGCGGATGATCACTGAGRTGAGAGTTGAGAGTTCAGCACTTTGGAGG 7623

QY 1936 ACCCGCTGTCTCAAAATGC--AAAAATTAGCCAGGCGTGTGACAGGACCTGTAT 1993
DB 7622 ACCCGCTGTCTCAAAATGC--AAAAATTAGCCAGGCGTGTGACAGGACCTGTAT 7563

QY 1994 CCCAGCTACTTGGAGGCTGAGGAGAGATGCTTGAACCCAGAGAGTGAAGTGTGA 2053
DB 7562 CCCAGCTACTTGGAGGCTGAGGAGAGATGCTTGAACCCAGAGAGTGAAGTGTGA 7503

QY 2054 GTAAGCTGAGATGTCGCTGTGACCTCAAGCTGCGGCAAGAGAGAGCTGTCTCA 2113
DB 7502 GTAGCGTGWATCRYGCATGTGCACTCCAGCTCGGAAACAAGAGAGCTCCGTCTCA 7443

QY 2114 GAAAAAAAAAAAAA 2129
DB 7442 AAAAAATATATATA 7427

RESULT 13
US-09-370-625A-8/c
; Sequence 8, Application US/09370625A
; Patent No. 6600032
; GENERAL INFORMATION:
; APPLICANT: Ford, John
```

```

; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/09/370, 625A
; CURRENT FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350, 836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273, 447
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1) .. (288)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1281) .. (1580)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1820) .. (1855)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (2467) .. (2555)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (2863) .. (2942)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (3889) .. (3950)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (4894) .. (4995)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (5847) .. (5987)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (6966) .. (7138)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (8556) .. (9365)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3409)
; OTHER INFORMATION: n = a or g or t or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9214)
; OTHER INFORMATION: n = a or g or t or c
; NAME/KEY: misc_feature
; LOCATION: (9303)
; OTHER INFORMATION: n = a or g or t or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9311)
; OTHER INFORMATION: n = a or g or t or c
US-09-370-625A-8

Query Match          9.4%; Score 265.6; DB 4; Length 9365;
Best Local Similarity 79.0%; Pred. No. 3.8e-49;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1765 CCAGCCTGGGCAACATACAGATCCTGTCTCTCAAAAAAAAAAAAAAAAAAGAAAG-1823
DB 7802 CCAGCATGGGTGACAGAGCAAGACCTGTCTCAAAAAAAAAAAAAAAAAAGAAAG 7743
```


Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY	1765	CCACCCTGGGCAACATATAACAAGATCCCTGCTCTCAAAAAAAAAAAAAAAAAAGAAAG-	1823
Db	11162	CCAGCATGGGTGACAGACCAAGACCCCTGTTCAAAAAAAAAAAAAAAAAAGAAAAAAG	110303
QY	1824	-----AGAGAGGCGGGCGGTGAGTCAAGCCTGTAAATCCAGCACTTTGGGAGG	1875
Db	11102	AAMGAMCAAAAAAGGCTRGMGGTGGTCTCAACTRTAATYVAGCACTTTGGGAGG	110433
QY	1876	CCGAGCCGGGGCGGATCACTGTGTCAAGAGTTTGAACCAAGCTTGCCAACTGGCAAA	1935
Db	11042	CYGGGGTGGGGATCACCTGAGRTCAAGATTYRGAAMVAGCTGGCCAAATGGTGA	109833
QY	1936	ACCCGCTGTACTCAAAATGC--AAAAATTAGCCAGCGCGTGTACAGCAACCTGTAT	1993
Db	10982	ACCCGCTCTACTAAATAATCAAAAAAATTAGCCRGCGTGTGGCGCRVGCCTGTANT	109233
QY	1994	CCAGCTACTTGGGAGGCTGAGGCAAGAGATCGCTTGAACCCAGAGGTGAGGTTTGA	2053
Db	10922	CCGAGTACTTGGGAGGCTGAGGCAAGATCGCTTGAACCCAGAGGTGAGGTTTGA	108633
QY	2054	GTAGGCTGAGATCGTGGCGCTTGCACTCCAGCTGGGCGACAGAGCAAGATCTGTCTCA	2113
Db	10862	GTAGGCTGAGATCGGCAATGCACTCCAGCTGGGGAACAGAGCRAGATCTCGTCTCA	108033
QY	2114	GAIAAAAAAAAAAAAAA 2129	
Db	10802	AAAAAAAAATATATATAA 10787	

```

Search completed: March 7, 2004, 11:59:47
Job time : 222 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 09:57:52 ; Search time 964 Seconds

(without alignments)
10691.539 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818
Sequence: 1 agagaaccatcatcatgta.....aaaaaaaaaaaaaaaaa 2818

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2421054 seqs, 1828716029 residues

Total number of hits satisfying chosen parameters: 4842108

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2818	100.0	2818	9	US-09-966-880A-7
2	2174.6	77.2	6564	9	US-09-966-880A-10
3	2174.6	77.2	11204	9	US-09-966-880A-35
4	2172	77.1	2172	9	US-09-966-880A-15
5	603.4	21.4	2440	9	US-09-966-880A-1
6	295.2	10.5	13862	10	US-09-764-891-5477
7	295.2	10.5	13862	10	US-09-764-891-10204
8	295.2	10.5	13862	14	US-10-205-428-10204
9	292.6	10.4	7809	10	US-09-764-891-6094
10	284.4	10.1	32349	10	US-09-764-891-7364
11	281.4	10.0	65608	9	US-09-962-436-232
12	281.4	10.0	65608	9	US-09-962-832-119
13	281.4	10.0	65608	9	US-09-954-531-180
14	280.2	9.9	11821	9	US-09-764-877-2857
15	280.2	9.9	11821	15	US-10-242-515-2857

16	280	9.9	15297	13	US-10-003-235-3	Sequence 3, Appli
17	278.2	9.9	32204	10	US-09-764-891-8537	Sequence 8537, Ap
18	277.4	9.8	30000	14	US-10-262-552-33	Sequence 33, Appl
19	276.6	9.8	37940	15	US-10-348-073A-12	Sequence 12, Appl
20	276.6	9.8	38000	15	US-10-175-627-11	Sequence 11, Appl
21	275.4	9.8	3965	14	US-09-764-855-227	Sequence 227, App
22	275.4	9.8	3965	14	US-10-072-349-227	Sequence 227, App
23	275	9.8	6482	14	US-10-180-410-6	Sequence 6, Appli
24	275	9.8	20050	14	US-10-180-410-5	Sequence 5, Appli
25	274.8	9.8	15362	9	US-09-764-877-2856	Sequence 2856, Ap
26	274.8	9.8	15362	15	US-10-242-515-2856	Sequence 2856, Ap
27	274.2	9.7	11811	10	US-09-728-552-7	Sequence 7, Appli
28	274.2	9.7	176001	16	US-10-210-556-27	Sequence 27, Appli
29	274.2	9.7	186739	16	US-10-210-556-19	Sequence 19, Appli
30	274	9.7	4963	9	US-09-764-877-2903	Sequence 2903, Ap
31	274	9.7	4963	15	US-10-242-515-2903	Sequence 2903, Ap
32	273.6	9.7	174424	10	US-09-967-768A-134	Sequence 314, App
33	273.6	9.7	174424	10	US-09-960-706-969	Sequence 969, App
34	273	9.7	2401	15	US-10-027-632-102635	Sequence 102635, Ap
35	273	9.7	26657	9	US-09-810-673A-3	Sequence 3, Appli
36	272	9.7	17491	14	US-10-395-242-3	Sequence 3, Appli
37	272	9.7	17491	14	US-10-017-161-1995	Sequence 1995, Ap
38	271	9.6	271	9	US-09-966-880A-13	Sequence 13, Appl
39	270.8	9.6	670	15	US-10-027-632-201375	Sequence 201375, Ap
40	270.8	9.6	1061	14	US-10-277-161-66	Sequence 66, Appl
41	270.8	9.6	107820	10	US-09-792-616-1	Sequence 1, Appli
42	270.2	9.6	17286	9	US-09-764-877-3234	Sequence 3234, Ap
43	270.2	9.6	17286	15	US-10-242-515-3334	Sequence 3234, Ap
44	269.8	9.6	43950	13	US-10-060-332-3	Sequence 3, Appli
45	269.8	9.6	43950	14	US-10-339-657-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-966-880A-7
Sequence 7, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Hoejio, Taseuku
FILE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09966, 880A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: PCT/JP00/01918
PRIOR FILING DATE: 2000-03-28
PRIOR APPLICATION NUMBER: JP 11-371382
PRIOR FILING DATE: 1999-12-27
PRIOR APPLICATION NUMBER: JP 11-178999
PRIOR FILING DATE: 1999-06-24
PRIOR APPLICATION NUMBER: JP 11-87192
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 2818
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (80)...(673)
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(79)
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (677)...(2818)
US-09-966-880A-7
Query Match 100.0%; Score 2818; DB 9; Length 2818;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAACCATCATTAATTAAGTAGAGATTTTCTGGCCGTGAGACTTGCAGGAGGCGAAGA 60
Db 1 AGAGAACCATCATTAATTAAGTAGAGATTTTCTGGCCGTGAGACTTGCAGGAGGCGAAGA 60
QY 61 AGACACTGTGACACCCCTATGACAGCCCTTGAAGAACCCGAGGAAGTTTCTTTACCA 120
Db 61 AGACACTGTGACACCCCTATGACAGCCCTTGAAGAACCCGAGGAAGTTTCTTTACCA 120
QY 121 ATTCAAAAATGTCGCGCTGAGGCTAAGGGTGGCGCTGAGACCTAAGCTGTGCTAGAGTGA 180
Db 121 ATTCAAAAATGTCGCGCTGAGGCTAAGGGTGGCGCTGAGACCTAAGCTGTGCTAGAGTGA 180
QY 181 GAGGCGTGAAGTGTCTATCTCTTTCTACTGACCTTGGTATCTTGGCAATPAGAACGG 240
Db 181 GAGGCGTGAAGTGTCTATCTCTTTCTACTGACCTTGGTATCTTGGCAATPAGAACGG 240
QY 241 CTGGCACTGTGGAATGTCTCTTCCCTCCGCTACATCTGGAGCTGGAGACTTGAACCTTGGCCG 300
Db 241 CTGGCACTGTGGAATGTCTCTTCCCTCCGCTACATCTGGAGCTGGAGACTTGAACCTTGGCCG 300
QY 301 CTGGTACCGGCTACCGCTGCTTCACTCTGAGAGCCCTGCTAGACTGTGACCCGAGCATGT 360
Db 301 CTGGTACCGGCTACCGCTGCTTCACTCTGAGAGCCCTGCTAGACTGTGACCCGAGCATGT 360
QY 361 GGGCGACTTCTGAGAGGAGACCCCAACTCACTGAGAGANTTTCACCGCGGCTCTA 420
Db 361 GGGCGACTTCTGAGAGGAGACCCCAACTCACTGAGAGANTTTCACCGCGGCTCTA 420
QY 421 CTCTGTGAGAGCCGCAAGGCTGAGCCCGAGGCGCTGCGCGGCTGACCGCGCGGAGT 480
Db 421 CTCTGTGAGAGCCGCAAGGCTGAGCCCGAGGCGCTGCGCGGCTGACCGCGCGGAGT 480
QY 481 GCAAAATGCGCATATGACCTTCAAGATTAATTTTCTGCTGAGAAATCTTTGTAGAAAA 540
Db 481 GCAAAATGCGCATATGACCTTCAAGATTAATTTTCTGCTGAGAAATCTTTGTAGAAAA 540
QY 541 CCATGAAGAAGACTTTCAAAGCTGGGAGGGGCTGATGAATAATTCAGTCTGCTCCAG 600
Db 541 CCATGAAGAAGACTTTCAAAGCTGGGAGGGGCTGATGAATAATTCAGTCTGCTCCAG 600
QY 601 ACAAGCTTGGCGGATCTTTGCCCCGTATGAGGTTGATGACTTACGAGACGCAATTGG 660
Db 601 ACAAGCTTGGCGGATCTTTGCCCCGTATGAGGTTGATGACTTACGAGACGCAATTGG 660
QY 661 TACTTTGGGACTTTGATGACACTTCCAGGAATGTCAACAGATGAATATCTCTGCTG 720
Db 661 TACTTTGGGACTTTGATGACACTTCCAGGAATGTCAACAGATGAATATCTCTGCTG 720
QY 721 AAGCAGTGTGATAAAAAGAGTCTTCAAGTCTTCTGTTTTATCTCAACTCTGAC 780
Db 721 AAGCAGTGTGATAAAAAGAGTCTTCAAGTCTTCTGTTTTATCTCAACTCTGAC 780
QY 781 TTTCTTGAAGTTTACGAAAAAATATTTATATAGACTCTTTTAAAAAGATCTATGCTTG 840
Db 781 TTTCTTGAAGTTTACGAAAAAATATTTATATAGACTCTTTTAAAAAGATCTATGCTTG 840
QY 841 AAAATAGAGAAAGAACAGAGTCTGGCCAGAGGACGCTGCAATTTGTCAAGTTTGAAT 900
Db 841 AAAATAGAGAAAGAACAGAGTCTGGCCAGAGGACGCTGCAATTTGTCAAGTTTGAAT 900
QY 901 GCAACATTTTCCCTTCTGGAATTAACAGACTGCGAGACCTGGAGGACTCTTAAGTGT 960
Db 901 GCAACATTTTCCCTTCTGGAATTAACAGACTGCGAGACCTGGAGGACTCTTAAGTGT 960
QY 961 CAAAGTTTCTATGACTTTTATGATGAGAGAGAGAGAGTATCTTAAAAAGCATG 1020
Db 961 CAAAGTTTCTATGACTTTTATGATGAGAGAGAGAGTATCTTAAAAAGCATG 1020
QY 1021 GTGAGAGGATCAATGTTTTATATCAACATCTTTATTAATTTGATTCATTTGAGTTAAC 1080
Db 1021 GTGAGAGGATCAATGTTTTATATCAACATCTTTATTAATTTGATTCATTTGAGTTAAC 1080

QY 1081 AGTGCTGTATGATTAATTTTCTATTTCTTTTCCCTTGAAGCTTTACTTTCAAGTAACAC 1140
Db 1081 AGTGCTGTATGATTAATTTTCTATTTCTTTTCCCTTGAAGCTTTACTTTCAAGTAACAC 1140
QY 1141 AAACCTTTCATGAGGCGATGATCTATAGAACCTCCCTAATGAGAGATCTGGGTGATTGT 1200
Db 1141 AAACCTTTCATGAGGCGATGATCTATAGAACCTCCCTAATGAGAGATCTGGGTGATTGT 1200
QY 1201 GACCCAAACCATCTCTCAAGCATTAATATCAATCATGCGCTGATGTTTTAAATCAG 1260
Db 1201 GACCCAAACCATCTCTCAAGCATTAATATCAATCATGCGCTGATGTTTTAAATCAG 1260
QY 1261 CAGAGCATGTTTTTATGTTTGTACAAAGAAAGATTGTTATGGGTGGGATGAGGTATA 1320
Db 1261 CAGAGCATGTTTTTATGTTTGTACAAAGAAAGATTGTTATGGGTGGGATGAGGTATA 1320
QY 1321 GACCATGATGTGACCTTGAAGTACTTTAATPAGAACTTAAATGGGCGAGAGGAC 1380
Db 1321 GACCATGATGTGACCTTGAAGTACTTTAATPAGAACTTAAATGGGCGAGAGGAC 1380
QY 1381 TGTGAACAAGACACCTAATATATGAGTGTCTGATGAGCAAACTCTTGAAGACG 1440
Db 1381 TGTGAACAAGACACCTAATATATGAGTGTCTGATGAGCAAACTCTTGAAGACG 1440
QY 1441 AAACCTTTTAAAGAACTCCCTAATTAATTAAGAAACCCCACTTCAATATCATATTA 1500
Db 1441 AAACCTTTTAAAGAACTCCCTAATTAATTAAGAAACCCCACTTCAATATCATATTA 1500
QY 1501 GCAAACTTGAAGAAAGATTGCTTGAATGTGGGAGAGAAAAATCTATTTGCTCTCGT 1560
Db 1501 GCAAACTTGAAGAAAGATTGCTTGAATGTGGGAGAGAAAAATCTATTTGCTCTCGT 1560
QY 1561 GGGTCTGTATCTCAGAAATGCAATCAGGTCAAGTGTGCTACATTTTGTATGTGTGT 1620
Db 1561 GGGTCTGTATCTCAGAAATGCAATCAGGTCAAGTGTGCTACATTTTGTATGTGTGT 1620
QY 1621 GATGCTTCTCCAAAGATTAATTAATTAAGAGATTTGACAAAACAGATGATTA 1680
Db 1621 GATGCTTCTCCAAAGATTAATTAATTAAGAGATTTGACAAAACAGATGATTA 1680
QY 1681 AGCTGCAACCGTGGCACAGCTCATATGTTCTGAGCTTGGAGGCTTGGAGGAGGA 1740
Db 1681 AGCTGCAACCGTGGCACAGCTCATATGTTCTGAGCTTGGAGGCTTGGAGGAGGA 1740
QY 1741 TGGCTTGAACAAGTGTTCAGAGCCAGCTGGGCAACATTAAGATCTGTCTTCAA 1800
Db 1741 TGGCTTGAACAAGTGTTCAGAGCCAGCTGGGCAACATTAAGATCTGTCTTCAA 1800
QY 1801 AAAAAAAGAAAAAAGAAAGAGAGGAGCGGAGCTGTGAGTCAAGCTGTATTC 1860
Db 1801 AAAAAAAGAAAAAAGAAAGAGAGGAGCGGAGCTGTGAGTCAAGCTGTATTC 1860
QY 1861 CAGCACTTGGAGGCGGAGCCGAGCTCACTGTGTGCTCAGAGTTTGAACAAGCTT 1920
Db 1861 CAGCACTTGGAGGCGGAGCCGAGCTCACTGTGTGCTCAGAGTTTGAACAAGCTT 1920
QY 1921 GGGCAATGAGCAAAACCCGCTGTATCTCAAAATGCAAAATTAAGCAGGCGTGTAGC 1980
Db 1921 GGGCAATGAGCAAAACCCGCTGTATCTCAAAATGCAAAATTAAGCAGGCGTGTAGC 1980
QY 1981 AGGCACTGTATCCAGCTACTTGGAGGCTGAGGAGAGGAATGCTTGAACCAAGA 2040
Db 1981 AGGCACTGTATCCAGCTACTTGGAGGCTGAGGAGAGGAATGCTTGAACCAAGA 2040
QY 2041 GGTGAGGTTGCAAGTATGATGATGCTGCTGTGCACTCAGCCTTGGGCGACAAGACA 2100
Db 2041 GGTGAGGTTGCAAGTATGATGATGCTGCTGTGCACTCAGCCTTGGGCGACAAGACA 2100
QY 2101 AGACTTGTCTCAGAAAAAAGAAAAAAGAGAGAGAGAGAGAAAGAAATTTTG 2160
Db 2101 AGACTTGTCTCAGAAAAAAGAAAAAAGAGAGAGAGAGAGAAAGAAATTTTG 2160

```

QY 2161 GGAAGAGAGATGGGAGAGATTCAGAGAAATTTGCTTTATCCAGAAATGTAAAGGA 2220
DB 2161 GGAAGAGAGATGGGAGAGATTCAGAGAAATTTGCTTTATCCAGAAATGTAAAGGA 2220
QY 2221 GCCAATAGAGGATCCCTATTGTCCTTTGCTGTCATTTGCTCCCTAAACAATGCTCTT 2280
DB 2221 GCCAATAGAGGATCCCTATTGTCCTTTGCTGTCATTTGCTCCCTAAACAATGCTCTT 2280
QY 2281 GACAGTGAAGAAATATTCGAAATACCATATCCCTGCTGCTGCTTATTAACCTGCAACCT 2340
DB 2281 GACAGTGAAGAAATATTCGAAATACCATATCCCTGCTGCTGCTTATTAACCTGCAACCT 2340
QY 2341 TGCATATGAGATGAGAGATTCACAGAGAAATCTGAAATGCAACCTGCTTATTTAAAC 2400
DB 2341 TGCATATGAGATGAGAGATTCACAGAGAAATCTGAAATGCAACCTGCTTATTTAAAC 2400
QY 2401 TTATGTACATTAAGTTGTAAAGAGTTAAAGATTTGCTTACTTCAATGATTTATATAT 2460
DB 2401 TTATGTACATTAAGTTGTAAAGAGTTAAAGATTTGCTTACTTCAATGATTTATATAT 2460
QY 2461 TTATATATTTTTCGCTTAATGATTTTATTAACATGATTTCTTTCTGATATATGA 2520
DB 2461 TTATATATTTTTCGCTTAATGATTTTATTAACATGATTTCTTTCTGATATATGA 2520
QY 2521 AATGAGATCTCAAGCTTCATTAATTTATTAACCTTTGAATGATTTATTAACAACGAT 2580
DB 2521 AATGAGATCTCAAGCTTCATTAATTTATTAACCTTTGAATGATTTATTAACAACGAT 2580
QY 2581 GTAATGTATACATTCAGATGATGCTGCTACAGAGCAATTTCTCTGATTTTATTAAGT 2640
DB 2581 GTAATGTATACATTCAGATGATGCTGCTACAGAGCAATTTCTCTGATTTTATTAAGT 2640
QY 2641 TTATGACACAGCAATTTGCTTCTGCTCACTTCAATGATTTAAATTAATGATTAAT 2700
DB 2641 TTATGACACAGCAATTTGCTTCTGCTCACTTCAATGATTTAAATTAATGATTAAT 2700
QY 2701 TTGGAAGCTGTAGAGATTAATACCAATTAATTAATTAATTAATTAATTAATGAAGT 2760
DB 2701 TTGGAAGCTGTAGAGATTAATACCAATTAATTAATTAATTAATTAATTAATGAAGT 2760
QY 2761 TAAATATTAATAATCAGATGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 2818
DB 2761 TAAATATTAATAATCAGATGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 2818

```

RESULT 2

US-09-966-880a-10
 ; Sequence 10, Application US/09966880A
 ; Patent No. US2002016473A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Honjo, Taku
 ; APPLICANT: Muramatsu, Masamichi
 ; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
 ; FILE REFERENCE: 06501-088001
 ; CURRENT APPLICATION NUMBER: US/09/966, 880A
 ; PRIOR APPLICATION NUMBER: PCT/JP00/01918
 ; PRIOR FILING DATE: 2000-03-28
 ; PRIOR APPLICATION NUMBER: JP 11-371382
 ; PRIOR FILING DATE: 1999-12-27
 ; PRIOR APPLICATION NUMBER: JP 11-178999
 ; PRIOR FILING DATE: 1999-06-24
 ; PRIOR APPLICATION NUMBER: JP 11-87192
 ; PRIOR FILING DATE: 1999-03-29
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10
 ; LENGTH: 6564
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-966-880A-10

Query Match 77.2%; Score 2174.6; DB 9; Length 6564;

```

Beat Local Similarity 99.8%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 617 CTTTGCCCTGATAGAGTTGATGATCTTACGAGACCAATTTGATCTTTGGGACTTTGA 676
DB 3735 CTACAGCCCTGATAGAGTTGATGATCTTACGAGACCAATTTGATCTTTGGGACTTTGA 3794
QY 677 TAGCACTTCCAGAAATGTCACACAGATGAATATCTTCTGAGAGACGTGATAAA 726
DB 3795 TAGCACTTCCAGAAATGTCACACAGATGAATATCTTCTGAGAGACGTGATAAA 3854
QY 737 AACGCTCTCAAGCTCTCTGTTTTTATTTCTCACTCTCACTTTCTTAAGTTTACA 796
DB 3855 AACGCTCTCAAGCTCTCTGTTTTTATTTCTCACTCTCACTTTCTTAAGTTTACA 3914
QY 797 GAAAAATATTTATATATGCACTCTTAAAGATCTATGCTGTAATAGAGAGAC 856
DB 3915 GAAAAATATTTATATATGCACTCTTAAAGATCTATGCTGTAATAGAGAGAC 3974
QY 857 ACAGGCTGGCCAGAGACGTGTCATTTGATGATGATGATGATGATGATGATGATGATGAT 916
DB 3975 ACAGGCTGGCCAGAGACGTGTCATTTGATGATGATGATGATGATGATGATGATGATGAT 4034
QY 917 CTGGGATTAACAGAACTGACAGACCTGGGACATCTTAAAGTCAACGTTTTCTATGA 976
DB 4035 CTGGGATTAACAGAACTGACAGACCTGGGACATCTTAAAGTCAACGTTTTCTATGA 4094
QY 977 CTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1036
DB 4095 CTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4154
QY 1037 TTTTATATCAACATCTTTATTTATTTGATTTATTTGATTTATTTGATTTATTTGATTTAT 1096
DB 4155 TTTTATATCAACATCTTTATTTATTTGATTTATTTGATTTATTTGATTTATTTGATTTAT 4214
QY 1097 GATTTTCTATCTTTTCCCTTGAAGTTTACTTCAAGTAAACAACCTTCCATCAGG 1156
DB 4215 GATTTTCTATCTTTTCCCTTGAAGTTTACTTCAAGTAAACAACCTTCCATCAGG 4274
QY 1157 CCATGATCTATAGACCTCTCTATAGAGATCTGGGTGATTTGATGCCCAACATCTC 1216
DB 4275 CCATGATCTATAGACCTCTCTATAGAGATCTGGGTGATTTGATGCCCAACATCTC 4334
QY 1217 TCCAAAGCATTAATATCATCATGAGCTGATGTTTATTCAGAGAGCATGTTTAT 1276
DB 4335 TCCAAAGCATTAATATCATCATGAGCTGATGTTTATTCAGAGAGCATGTTTAT 4394
QY 1277 TGTTTGACAAAAGAGATTTGATGGTGGGATGAGGATATAGACATGATGTCAC 1336
DB 4395 TGTTTGACAAAAGAGATTTGATGGTGGGATGAGGATATAGACATGATGTCAC 4454
QY 1337 CTTCAGACTTCTTAATTAAGATCTTAAATGGGAGAGAGACCTGTAACAACACCC 1396
DB 4455 CTTCAGACTTCTTAATTAAGATCTTAAATGGGAGAGAGACCTGTAACAACACCC 4514
QY 1397 TAAATAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1456
DB 4515 TAAATAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4574
QY 1457 GTCCCTATTTAGAAACACCAACCACTTCAATCATATAGAGAAACAATTTGAAG 1516
DB 4575 GTCCCTATTTAGAAACACCAACCACTTCAATCATATAGAGAAACAATTTGAAG 4634
QY 1517 AAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1576
DB 4635 AAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4694
QY 1577 GAAATGCAATCAGGTCAGGTTTCTATCTTTGATGATGATGATGATGATGATGATGATGATGAT 1636
DB 4695 GAAATGCAATCAGGTCAGGTTTCTATCTTTGATGATGATGATGATGATGATGATGATGATGAT 4754
QY 1637 GTATATTAATATATAGAGATTTGACAAAACAATGATTAAGCTGGAACGCTGGC 1696

```

Db 4755 GTATATTACTATATAAGAGATTGTGACAAAACAGATGATTAAGCTGCGAACCGTGGC 4814
Qy 1697 ACACGCTCATAGTTCTTAGCTGCTGGAGGTTGAGAGGAGGATGGCTTGAACACAGGT 1756
Db 4815 ACAAGCTCATAGTTCTTAGCTGCTGGAGGTTGAGAGGAGGATGGCTTGAACACAGGT 4874
Qy 1757 GTTCAAGGCGACGCTGGGCGAACAATAACAGATCTCTCTCAAAAAAAAAAAAAAAAAAAAA 1816
Db 4875 GTTCAAGGCGACGCTGGGCGAACAATAACAGATCTCTCTCAAAAAAAAAAAAAAAAAAAAA 4934
Qy 1817 AAGAAAGAGAGAGGCGCGGCGTGTGCTCACGCTGTATCCAGCACTTTGGAGGCG 1876
Db 4935 AAGAAAGAGAGAGGCGCGGCGTGTGCTCACGCTGTATCCAGCACTTTGGAGGCG 4994
Qy 1877 CGAGCCGCGGCGATCACTGTGTGTCAGAGATTGAGACCAAGCTGCGCAACATGCGAAA 1936
Db 4995 CGAGCCGCGGCGATCACTGTGTGTCAGAGATTGAGACCAAGCTGCGCAACATGCGAAA 5054
Qy 1937 CCCCCTGTGTACTTCAAAATGCAAAATTAACCGAGGCTGGTAGCAGGCACTGTATCCC 1996
Db 5055 CCCCCTGTGTACTTCAAAATGCAAAATTAACCGAGGCTGGTAGCAGGCACTGTATCCC 5114
Qy 1997 AGCTACTTGGAGGCTGAGGCGAGAGATGCTTGAACCCAGAGGTTGGAGTTGAGTA 2056
Db 5115 AGCTACTTGGAGGCTGAGGCGAGAGATGCTTGAACCCAGAGGTTGGAGTTGAGTA 5174
Qy 2057 AGCTGAGATCGTGCCTGTGCACTCCAGCTGCGCGACCAAGAGCAACTGTCTCGAGA 2116
Db 5175 AGCTGAGATCGTGCCTGTGCACTCCAGCTGCGCGACCAAGAGCAACTGTCTCGAGA 5234
Qy 2117 AAAAAAAAAAAAAAAAAAG 2176
Db 5235 AAAAAAAAAAAAAAAAAAG 5294
Qy 2177 AAGATGTCAGAGAAATTTGCTTTATCCCAACAAATGTAAGAGGCAATTAAGGATCCC 2236
Db 5295 AAGATGTCAGAGAAATTTGCTTTATCCCAACAAATGTAAGAGGCAATTAAGGATCCC 5354
Qy 2237 TATTTGCTCTTTTGGTGTCTATTTGCTTAACTGCTTGAAGAGAGAGAGAGAGAGAG 2296
Db 5355 TATTTGCTCTTTTGGTGTCTATTTGCTTAACTGCTTGAAGAGAGAGAGAGAGAGAG 5414
Qy 2297 TTCAAGTAACCAATATCCCTGTGCGTATTATTAACCTGAGCAACCTTGCAATGAAGATGAC 2356
Db 5415 TTCAAGTAACCAATATCCCTGTGCGTATTATTAACCTGAGCAACCTTGCAATGAAGATGAC 5474
Qy 2357 AGATCCACAGAGAAACTTGAATGACACAGCTCTTATTTAATCTTATGTACATAAGT 2416
Db 5475 AGATCCACAGAGAAACTTGAATGACACAGCTCTTATTTAATCTTATGTACATAAGT 5534
Qy 2417 TGTAAAGAGTTAAATTTGTTACTGATGATTCATTTATTTATTTATTTATTTTGGCT 2476
Db 5535 TGTAAAGAGTTAAATTTGTTACTGATGATTCATTTATTTATTTATTTATTTTGGCT 5594
Qy 2477 CTAATGATTTTATTAACATGATTTCCCTTTCTGATATTAATGAATGAGAGCTCAAGC 2536
Db 5595 CTAATGATTTTATTAACATGATTTCCCTTTCTGATATTAATGAATGAGAGCTCAAGC 5654
Qy 2537 TTCAATTAATTTATTAACATGATTTCTTAATTAACACGATGATGATTTGAACATTCG 2596
Db 5655 TTCAATTAATTTATTAACATGATTTCTTAATTAACACGATGATGATTTGAACATTCG 5714
Qy 2597 AGTAAGTGTGTAAGAGCAATTTCTGATTTTATTTATTTATTTATTTATTTATTTATTT 2656
Db 5715 AGTAAGTGTGTAAGAGCAATTTCTGATTTTATTTATTTATTTATTTATTTATTTATTT 5774
Qy 2657 TGCTTCTGCTCACTTTCAATCAAGTAAATTAATTAATTAATTTTGAAGCTGTGAAG 2716
Db 5775 TGCTTCTGCTCACTTTCAATCAAGTAAATTAATTAATTAATTTTGAAGCTGTGAAG 5834
Qy 2717 ATAAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2776
Db 5835 ATAAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5894

Qy 2777 TATGATGAATTAACCTTGAAA 2797
Db 5895 TATGATGAATTAACCTTGAGA 5915

RESULT 3
US-09-966-880a-35
; Sequence 35, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966, 880A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 11204
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880a-35

Query Match 77.2%; Score 2174.6; DB 9; Length 11204;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 617 CTTTGGCCCTGTATGAGGTTGATGACTTACGAGACCACTTGTGACTTTGGAGCTTTGA 676
Db 8950 CTAGAGCCCTGTATGAGGTTGATGACTTACGAGACCACTTGTGACTTTGGAGCTTTGA 9009
Qy 677 TAGCACTTCCAGGAATGTCAACACAGATGAATATCTCTGCTGAGACAGCTGGATATAA 736
Db 9010 TAGCACTTCCAGGAATGTCAACACAGATGAATATCTCTGCTGAGACAGCTGGATATAA 9069
Qy 737 AACAGCTCTCAATCTCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 796
Db 9070 AACAGCTCTCAATCTCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 9129
Qy 797 GAAAAAATATTTATATAGCACTCTTTAAAGATCTATGCTTGAATAATAGAGAGAGAC 856
Db 9130 GAAAAAATATTTATATAGCACTCTTTAAAGATCTATGCTTGAATAATAGAGAGAGAC 9189
Qy 857 ACAGGCTGCGCAGAGAGCTGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCTGCT 916
Db 9190 ACAGGCTGCGCAGAGAGCTGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCTGCT 9249
Qy 917 CTGGGAATTAACAGACTGCGAGAGCTGGAGACCTCTTAAGTGTCAACGTTTTTCTATGA 976
Db 9250 CTGGGAATTAACAGACTGCGAGAGCTGGAGACCTCTTAAGTGTCAACGTTTTTCTATGA 9309
Qy 977 CTTTAAAGTAG 1036
Db 9310 CTTTAAAGTAG 9369
Qy 1037 TTTTATATCAACATCTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1096
Db 9370 TTTTATATCAACATCTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 9429
Qy 1097 GATTTTCTATTTCTTTCCCTGAGCTTACTTTCAAGTAAACACAACTTTTCATCAGG 1156
Db 9430 GATTTTCTATTTCTTTCCCTGAGCTTACTTTCAAGTAAACACAACTTTTCATCAGG 9489

QY	1157	CCAGTACCTTATAGAACCTCCCTAATAGACAGATCTGGGTGATTGGACCCCAACCATTC	1216
Db	9490	CCATGATCTATAGACCTCTCTAATAGAGATATCTGGGTATTTGTACCCCAACCATCTTC	9549
QY	1217	TCCAAGCATTAAATATCCAAATCATGCGCTGTATGTTTAAATCAGCAGAAAGATGTTTAA	1276
Db	9550	TCCAAGCATTAAATATCCAAATCATGCGCTGTATGTTTAAATCAGCAGAAAGATGTTTAA	9609
QY	1277	TGTTTGTACAAAAGAAATTTGTTATGGGTGGGATGAGGTATATGACATGATGGTCAC	1336
Db	9610	TGTTTGTACAAAAGAAATTTGTTATGGGTGGGATGAGGTATATGACATGATGGTCAC	9669
QY	1337	CTTCAGCTACTTAAATAAAGATCTTAAATGGGCGAGAGACGTGTGAACAAGACACC	1396
Db	9670	CTTCAGCTACTTAAATAAAGATCTTAAATGGGCGAGAGACGTGTGAACAAGACACC	9729
QY	1397	TAATPATGGGTGATGTCTGATAGTACAAATCTTCTGGAACGCAAACTCTTTAAGAA	1456
Db	9730	TAATPATGGGTGATGTCTGATAGTACAAATCTTCTGGAACGCAAACTCTTTAAGAA	9789
QY	1457	GTCCCTAATTTAGAAACCCACAACTTCAATATCAATATTAGCAACAAATGGAAG	1516
Db	9790	GTCCCTAATTTAGAAACCCACAACTTCAATATCAATATTAGCAACAAATGGAAG	9849
QY	1517	AAGTTGCTGTGAATGTTGGGAGAGAAATCTATGTCCTCGTGGGTCTCTTCACTCA	1576
Db	9850	AAGTTGCTGTGAATGTTGGGAGAGAAATCTATGTCCTCGTGGGTCTCTTCACTCA	9909
QY	1577	GAAATGCAATACAGGTCAAGTTTGCTATCATTTTGTATGTGTGATGCTTCCCAAG	1636
Db	9910	GAAATGCAATACAGGTCAAGTTTGCTATCATTTTGTATGTGTGATGCTTCCCAAG	9969
QY	1637	GTAATATTACTATATTAAGAGAGTTGTGCAAAAACAGATGATTAAGCTGCAACCGTGC	1696
Db	9970	GTAATATTACTATATTAAGAGAGTTGTGCAAAAACAGATGATTAAGCTGCAACCGTGC	10022
QY	1697	ACAGGCTCATAGTTCTTAGCTGCTTGGGAGGTGTGAGAGGAGATGGCTTGAACAAGT	1756
Db	10030	ACAGGCTCATAGTTCTTAGCTGCTTGGGAGGTGTGAGAGGAGATGGCTTGAACAAGT	10083
QY	1757	GTTCAAGGCGACGCTGGGCAACATPAACAATCTGTCTCTCAAAAAAAAAAAAAAAAAA	1816
Db	10090	GTTCAAGGCGACGCTGGGCAACATPAACAATCTGTCTCTCTCAAAAAAAAAAAAAAAAAA	10144
QY	1817	AAGAAAGAGAGGCGCGGCGTGTGTGCTCACGCGTGTATCCAGCACTTTGGAGGC	1876
Db	10150	AAGAAAGAGAGGCGCGGCGTGTGTGCTCACGCGTGTATCCAGCACTTTGGAGGC	10205
QY	1877	CGAGCCGGGCGGATCACCTGTGGTCAAGAGTTTGACACACGCTGGCCAAATGCGAATA	1936
Db	10210	CGAGCCGGGCGGATCACCTGTGGTCAAGAGTTTGACACACGCTGGCCAAATGCGAATA	10263
QY	1937	CCCGGCTGTACTCAAAAATGCAAAAATTAGCCAGGGGTGTATCAAGCACTGTATCCC	1996
Db	10270	CCCGGCTGTACTCAAAAATGCAAAAATTAGCCAGGGGTGTATCAAGCACTGTATCCC	10322
QY	1997	AGTACTTTGGAGGCTGAGCAGAGAGATGCTTGAACCAAGAGGTGAGATTTGACATTA	2056
Db	10330	AGTACTTTGGAGGCTGAGCAGAGAGATGCTTGAACCAAGAGGTGAGATTTGACATTA	10388
QY	2057	AGCTGAATGTGTCCTGTTCATCTCCAGCTGGGCGACAGAGCAATCTGTCTCAGAA	2116
Db	10390	AGCTGAATGTGTCCTGTTCATCTCCAGCTGGGCGACAGAGCAATCTGTCTCAGAA	10443
QY	2117	AAAAAAAAAAAAAAAAAGAGAGAGAGAGAAAGAGAACTTTTGGAGAGAAAGATGGGG	2176
Db	10450	AAAAAAAAAAAAAAAAAGAGAGAGAGAGAAAGAGAACTTTTGGAGAGAAAGATGGGG	10503
QY	2177	AAGCATTGCAAGAAATTTGCTTTATCCAAACAATGTATAGAGCCATATAGGATCCC	2236
Db	10510	AAGCATTGCAAGAAATTTGCTTTATCCAAACAATGTATAGAGCCATATATAGGATCCC	10566
QY	2237	TATTTGTCTCTTTGGTGTCTAATTTGTCCCTAACACTGTCTTTGACATGAGAAAAATA	2296

Db	10570	TATTTCCTCTTTGGTGTCTATTGTCCTTAACACTGTCTTTGACGTGAAAAATA	106238
OY	2297	TTGAAATTAACCATATCCCTGTGCGGTATTAACCTAGCAACCCCTGCATGGAATGAGC	2356
Db	10630	TTGGAATTAACCATATCCCTGTGCGGTATTAACCTAGCAACCCCTGCATGGAATGAGC	10689
OY	2357	AGATCCACGAGAAACTTGAAATGCAACACTGTCTTATTTTAATCTTATGTACATAGTT	2416
Db	10690	AGATCCACGAGAAAACCTTGAAATGCAACAACCTGTCTTATTTTAATCTTATGTACATAGTT	10749
OY	2417	TGTAAAGAGTTAAAAATGTATCTTCATGATATCACTTAATATTTAATTTTGGCT	2476
Db	10750	TGTAAAGAGTTAAAAATGTATCTTCATGATATCACTTAATATTTAATTTTGGCT	10809
OY	2477	CTATGATTTTTTATTATACATGATTTCCCTTTCTGATATATGAAATGAGATCMAAGC	2536
Db	10810	CTATGATTTTTTATTATACATGATTTCCCTTTCTGATATATGAAATGAGATCMAAGC	10866
OY	2537	TTCTAAATTTATTAACCTTTAGAAATGATTTCTAATAACAAGTAGTAATTTGTAACCTGC	2596
Db	10870	TTCTAAATTTATTAACCTTTAGAAATGATTTCTAATAACAAGTAGTAATTTGTAACCTGC	10923
OY	2597	AGTATGATGCTACGAAGCCATTTCTTGATTTTATGTAACCTTTATGACAGCAATT	2656
Db	10930	AGTATGATGCTACGAAGCCATTTCTTGATTTTATGTAACCTTTATGACAGCAATT	10989
OY	2657	TGCTTCGCGCTCACTTCAATCACTTAATTAATGATAATATTTTGGAACTGTGAAG	2716
Db	10990	TGCTTCGCGCTCACTTCAATCACTTAATTAATGATAATATTTTGGAACTGTGAAG	11049
OY	2717	ATTAATAACCAATTAATAATTAATAAAGTATTAATGAAGCTTAATAATAAATATCG	2776
Db	11050	ATTAATAACCAATTAATAATTAATAAAGTATTAATGAAGCTTAATAATAAATATCG	11099
OY	2777	TATGATGGAATTAACCTGAAA 2797	
Db	11110	TATGATGGAATTAACCTGAGA 11130	

```

        623 CCCCTGTATGACGTTGATGACTTACGAGACGCACTTTGTCACCTTTGGGACCTTGATGACA 682
US-09-966-880A-15
; Sequence 15, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Matsumoto, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: FCT/JPO/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-176999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-87192
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 2172
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-966-880A-15
Query Match 77.1%; Score 2172; DB 9; Length 2172;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

Db 1 CCCCCGATGAGGTGATGACCTTACGAGACGATTTCTGACTTTGGGACCTTTGATAGCA 60
QY 683 CTTCGAGAGATGTCAACACGATGAAAATCTCTGCTGTAAGACAGTGATTAAGAAACAGT 742
Db 61 CTTCGAGAGATGTCAACACGATGAAAATCTCTGCTGTAAGACAGTGATTAAGAAACAGT 120
QY 743 CATTGAGATGCTCTGCTGCTTTTATTTCTGACCTTCACTTTCTTGAGTTTCAAGAAA 802
Db 121 CATTGAGATGCTCTGCTGCTTTTATTTCTGACCTTCACTTTCTTGAGTTTCAAGAAA 180
QY 803 ATATTATATACAGACTCTTTAAAGATCTATGCTTTGAAAATAGAGAGAACAGAGT 862
Db 181 ATATTATATACAGACTCTTTAAAGATCTATGCTTTGAAAATAGAGAGAACAGAGT 240
QY 863 CTGCGCAGGAGCTGCTGCTGCAATTTGTCAGTTTGAATGCAATTTGCTCCCTACTGCGA 922
Db 241 CTGCGCAGGAGCTGCTGCTGCAATTTGTCAGTTTGAATGCAATTTGCTCCCTACTGCGA 300
QY 923 ATTAACGAACTGACAGACCTGGGAGCATCTTAAGTCAAGCTTTTCTATGACTTTTA 982
Db 301 ATTAACGAACTGACAGACCTGGGAGCATCTTAAGTCAAGCTTTTCTATGACTTTTA 360
QY 983 GATGAGATGAGACAGAGGTAGATCTTAAGAAAGCATGGTGAAGGATCAATGTTTTTA 1042
Db 361 GATGAGATGAGACAGAGGTAGATCTTAAGAAAGCATGGTGAAGGATCAATGTTTTTA 420
QY 1043 TATCAACATCTTTATTTATTTATTTGATTCATTTGATTAACAGTGTGTTAGTAGATTT 1102
Db 421 TATCAACATCTTTATTTATTTATTTGATTCATTTGATTAACAGTGTGTTAGTAGATTT 480
QY 1103 TCTATTTCTTTCCCTTGACGTTTACTTTCAAGTACCAAACTCTTCATCAGGSCATGA 1162
Db 481 TCTATTTCTTTCCCTTGACGTTTACTTTCAAGTACCAAACTCTTCATCAGGSCATGA 540
QY 1163 TCTATGAGACCTCTCTATGAGAGTATCTGGGTGATTTGTAACCCCAACCATCTCCCAA 1222
Db 541 TCTATGAGACCTCTCTATGAGAGTATCTGGGTGATTTGTAACCCCAACCATCTCCCAA 600
QY 1223 GCATTATATTCATCATGAGCGCTGTATGTTTTAATCAGCAGAGAGCTTTTATGTTTG 1282
Db 601 GCATTATATTCATCATGAGCGCTGTATGTTTTAATCAGCAGAGAGCTTTTATGTTTG 660
QY 1283 TACAAAAGAAAGATTGTTATGGGTGGGATGAGAGTATAGACATGATGTCACCTTCA 1342
Db 661 TACAAAAGAAAGATTGTTATGGGTGGGATGAGAGTATAGACATGATGTCACCTTCA 720
QY 1343 GCTACTTTATTAAGAGATCTTAAGATGGGACAGAGACTGTGAACAACACCTTAATPA 1402
Db 721 GCTACTTTATTAAGAGATCTTAAGATGGGACAGAGACTGTGAACAACACCTTAATPA 780
QY 1403 TGGGTTGATGCTGTAAGTAGCAAACTCTGCGAAAAGCAAACTCTTTTAAGAGAGTCCCT 1462
Db 781 TGGGTTGATGCTGTAAGTAGCAAACTCTGCGAAAAGCAAACTCTTTTAAGAGAGTCCCT 840
QY 1463 AATTAGAAACACCAACCACTTCAATATCATATTAAGCAAACTTTGAAGAGAGTTG 1522
Db 841 AATTAGAAACACCAACCACTTCAATATCATATTAAGCAAACTTTGAAGAGAGTTG 900
QY 1523 CTGGAATGTTGGGAGAGAGAAATCTATTTGGCTCTGCTGGGTCTCTTCACTTCAAGAAATG 1582
Db 901 CTGGAATGTTGGGAGAGAGAAATCTATTTGGCTCTGCTGGGTCTCTTCACTTCAAGAAATG 960
QY 1583 CCAATCAGGTCAAGTTTGTCTACATTTTGATGTTGTGATGCTTCTCCCAAGGATAT 1642
Db 961 CCAATCAGGTCAAGTTTGTCTACATTTTGATGTTGTGATGCTTCTCCCAAGGATAT 1020
QY 1643 TAACTATATAAGAGAGTTGTGACAAAACAGATTAAGCTGCGAACCGTGGCACACGC 1702
Db 1021 TAACTATATAAGAGAGTTGTGACAAAACAGATTAAGCTGCGAACCGTGGCACACGC 1080
QY 1703 TCAATAGTTCTAGCTGCTTGGAGAGTTGAGAGAGGAGATGGCTTGAACAAGGTGTTCA 1762
Db 1081 TCAATAGTTCTAGCTGCTTGGAGAGTTGAGAGAGGAGATGGCTTGAACAAGGTGTTCA 1140

QY 1763 GGCCAGCTCTGGGCAATTAACAAAGATCTGTCTCTCAAAAAAAAAAAAAAAAAAAAAA 1822
Db 1141 GGCCAGCTCTGGGCAATTAACAAAGATCTGTCTCTCAAAAAAAAAAAAAAAAAAAAAA 1200
QY 1823 GAGAGAGGCGGGGCGGTGGTGTCTCAAGCTGTATCCGAGCACTTTGGAGGCGAGGC 1882
Db 1201 GAGAGAGGCGGGGCGGTGGTGTCTCAAGCTGTATCCGAGCACTTTGGAGGCGAGGC 1260
QY 1883 GGGGAGATCACTGTGGTCAGAGATTGAAGCCAGCTGGCCAAACATGGCAAAACCCCGT 1942
Db 1261 GGGGAGATCACTGTGGTCAGAGATTGAAGCCAGCTGGCCAAACATGGCAAAACCCCGT 1320
QY 1943 CTGTACTCAAAATGCAAAATTTAGCCAGGCGTGTGACAGGCACTGTATCCAGCTAC 2002
Db 1321 CTGTACTCAAAATGCAAAATTTAGCCAGGCGTGTGACAGGCACTGTATCCAGCTAC 1380
QY 2003 TTGGAGAGCTGAGGACAGAGAAATCGCTTGAACCCAGAGAGTGGAGGTTGACATAGCTGA 2062
Db 1381 TTGGAGAGCTGAGGACAGAGAAATCGCTTGAACCCAGAGAGTGGAGGTTGACATAGCTGA 1440
QY 2063 GATGTCCTGTCATCTCAAGCTGGGCGCAAGAGCAAGACTGTCTCAAGAAAAAAA 2122
Db 1441 GATGTCCTGTCATCTCAAGCTGGGCGCAAGAGCAAGACTGTCTCAAGAAAAAAA 1500
QY 2123 AAAAAAAAAAGAGAGAGAGAGAAAGAACATATTTGGAGAGAGATGGGAAAGCAT 2182
Db 1501 AAAAAAAAAAGAGAGAGAGAGAAAGAACATATTTGGAGAGAGATGGGAAAGCAT 1560
QY 2183 TGCAGAGAAATTTGCTTTATCCAAAGAAATGTAAGAGCAATTAAGGATCCCTATTGG 2242
Db 1561 TGCAGAGAAATTTGCTTTATCCAAAGAAATGTAAGAGCAATTAAGGATCCCTATTGG 1620
QY 2243 TCTCTTTGGTGTCTATTTGCTCCCAACAGCTCTTGAAGAGTGAAGAAATATTCAGA 2302
Db 1621 TCTCTTTGGTGTCTATTTGCTCCCAACAGCTCTTGAAGAGTGAAGAAATATTCAGA 1680
QY 2303 ATTAACATATCCCTGTGCTGCTTTATTAACCTAGCAACCCCTGCAATGAGATGACAGATCC 2362
Db 1681 ATTAACATATCCCTGTGCTGCTTTATTAACCTAGCAACCCCTGCAATGAGATGACAGATCC 1740
QY 2363 ACAGAGAAACCTGATGACACATGCTCTTATTTTAATCTTATGTACATTAAGTTGTA 2422
Db 1741 ACAGAGAAACCTGATGACACATGCTCTTATTTTAATCTTATGTACATTAAGTTGTA 1800
QY 2423 AGAGTTAAATTTGTAATCTGATGATGATTAATTAATTTTGGGTCTAAG 2482
Db 1801 AGAGTTAAATTTGTAATCTGATGATGATTAATTAATTTTGGGTCTAAG 1860
QY 2483 ATTTTATTAATGATGATTTCTTTCTGATATATTAAGATGAGTCTCAAGCTTCATA 2542
Db 1861 ATTTTATTAATGATGATTTCTTTCTGATATATTAAGATGAGTCTCAAGCTTCATA 1920
QY 2543 AATTATTAATTTAGAAATGATTTCTAATACACATGATGATTTGTAACATTCAGATAT 2602
Db 1921 AATTATTAATTTAGAAATGATTTCTAATACACATGATGATTTGTAACATTCAGATAT 1980
QY 2603 GGTGCTACGAAGCATTTCTGTTGATTTTATAGTAACTTTATGACAGCAAAATTTGCTTC 2662
Db 1981 GGTGCTACGAAGCATTTCTGTTGATTTTATAGTAACTTTATGACAGCAAAATTTGCTTC 2040
QY 2663 TGGCTCACTTCAATCAGTTAAATTAATGATTAATTAATTTTGGAGCTGGAAGTAAAA 2722
Db 2041 TGGCTCACTTCAATCAGTTAAATTAATGATTAATTAATTTTGGAGCTGGAAGTAAAA 2100
QY 2723 TACCAATTAATTAATTAATTAAGTATTAATTAAGTAAATTAATTAATTAATTAATTAAT 2782
Db 2101 TACCAATTAATTAATTAATTAAGTATTAATTAAGTAAATTAATTAATTAATTAATTAAT 2160
QY 2783 GGAATTAACCTG 2794
Db 2161 GGAATTAACCTG 2172

RESULT 5
US-09-966-880A-1
Sequence 1, Application US/09966880A
Patent No. US20020164743A1
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Muramatsu, Masamichi
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
FILE REFERENCE: 06501-088001
CURRENT APPLICATION NUMBER: US/09/966, 880A
CURRENT FILING DATE: 2001-09-28
PRIORITY APPLICATION NUMBER: PCT/JP00/01918
PRIORITY FILING DATE: 2000-03-28
PRIORITY APPLICATION NUMBER: JP 11-371382
PRIORITY FILING DATE: 1999-12-27
PRIORITY APPLICATION NUMBER: JP 11-178999
PRIORITY FILING DATE: 1999-06-24
PRIORITY APPLICATION NUMBER: JP 11-871992
PRIORITY FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2440
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (93)...(686)
FEATURE:
NAME/KEY: 5'UTR
LOCATION: (1)...(92)
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (690)...(2440)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(2440)
OTHER INFORMATION: n = A,T,C or G
US-09-966-880A-1

Query Match 21.4%; Score 603.4; DB 9; Length 2440;
Best Local Similarity 69.4%; Pred. No. 4e-118;
Matches 1006; Conservative 0; Mismatches 351; Indels 92; Gaps 10;

QY 4 GAACCATCATTAATGAGTGAATTTCTGGCCGAGACTTGACGAGAGGCAAG-AAG 62
DB 16 GAGGCAAGCTTGTGAGAGCAAGCTTGTGGCTTAAGACTTTGAGGAGTCAAGAAAG 75
QY 63 ACACTGTGACACACATATGACAGCCTCTTGATGACCGAGAAAGTTCTTTACAAAT 122
DB 76 TCAGCGTGAAGACCATATGACAGCCTCTGATGAGCAAAAGAGTTCTTTACAAAT 135
QY 123 TCAAAATGTCCGCTGAGGCTAAGGGTGGCGTGAAGACTTACCTGTGCTACGTAGTGAAGA 182
DB 136 TCAAAATGTCCGCTGAGGCTAAGGGTGAAGGACGAGATGAGACTACCTGTGCTAGTGAAGA 195
QY 183 GGGGTGACGTGTATACCTTTTCACTGACCTTTGTGTTATCTTGCAATAGAAACGGCT 242
DB 196 GGGAGATATGTGCACTCTGCTCACTGACTTGGCCACCTTGGCAACAAAGTGTGCT 255
QY 243 GCCACGTGAAATGCTCTTCTCCGCTACATCTGGAAGTGGAGCTGAGCCCTGGCGCT 302
DB 256 GCCACGTGAAATGCTCTTCTCCGCTACATCTGGAAGTGGAGCTGAGCCCTGGCGCT 315
QY 303 GCTACCGCGTCACTGTGTTCACTCTGAGCCCTGCTACGATGATGTGCCGACATGTGG 362
DB 316 GTTACCGCGTCACTGTGTTCACTCTGAGCCCGGTATGATGTGTGCCGCGACGTGG 375
QY 363 CCGACTTTCTGAGAGGAAACCCCAAGCTCAAGTGTGAGATCTTACCGGCGGCTCTACT 422
DB 376 CTGAGTTCTGAGATGAACCCCAAGCTCAAGCTGAGGATTTTCAACCGGCGGCTCTACT 435

QY 423 TCTGTGAGAACCGGAGGCTGAGCCCGAGGCGCTGCGCGCTGCAACCGCGCGGCTGC 482
DB 436 TCTGTGAAGACCGGAGGCTGAGCCCGAGGCGCTGCGCGCTGCAACCGCGCGGCTGC 495
QY 483 AATAGCCATCATGACCTTCAAGATTAATTTTACTGCTGGAATACCTTTGTGAAAAAC 542
DB 496 AGATCGGATATATACCTTTCAAGACTATTTTACTGCTGGAATACCTTTGTGAAAAATC 555
QY 543 ATGAAAGACTTTCAAGCCCTGGAGAGGCTGCAATGAAATTCAGTTGCTCTCCAGAC 602
DB 556 GTGAAAGACTTTCAAGCCCTGGAGAGGCTGCAATGAAATTCAGTTGCTCTCCAGAC 615
QY 603 AGCTTGGCGGATCTTTTGGCCCTGTATGAGGTTGATGACTTACGAGACGCAATTCGTA 662
DB 616 AACTTGGCGGATCTTTTGGCCCTGTATGAGGTTGATGACTTACGAGACGCAATTCGTA 675
QY 663 CTTTGGGACTTGTATGACACTTCCAGAAATGTACACAGATGAAATATCTGCTGAA 722
DB 676 TGTGGGATTTTGAAGCAACTCTGGAATGTACAGTGAATATTTCT--CTGAA 732
QY 723 GACAGTGATTAATAAAGCTTCAAGTCTTCTGTTTATTTCTTCAACTCTCACTT 782
DB 733 GAGACTGATTAATAAAGCTTCAAGTCTTCTGTTTATTTCTTCAACTCTCACTT 790
QY 783 TCTTAGATTTACGAAAAAATTTATATACACTCTTTAAAAAGATCTATGCTTGA 842
DB 791 TTATAGTGTAGGGGAAA-----TTATATGACTTTTAAAAAATACTTGAAGCTGAC 843
QY 843 AATGAGAGAAACACAGCTGTGCGCAGGACGCTGCAATTTGAGCAATTTGAATGC 902
DB 844 AGACCGCAGAGGAATGATGTAAC-----TGAAGCTGTGTGC 882
QY 903 AACATGTCCCTACT-GGGAAATPACAGAACTGAGAGCTGGAGAGATCTAAAGTGC 961
DB 883 AACATGTCCCTACT-GGGAAATPACAGAACTGAGAGCTGGAGAGATCTAAAGTGC 942
QY 962 AACGTTTTCTAGACTTTTATGATGATGAGAGACAGAAAGTATCTTAAAAAGCATGG 1021
DB 943 TTTTTCCTCAAGACAT-----GGAAGGCAAT 971
QY 1022 TGAAGATCAATATTTTATATCAACATCTTATTTATTTGATTCATTTGATPACA 1081
DB 972 GGAAGACACACACAGTTTGTATACCCACCTGTGTTCTTGATTCATTTGATTCACA 1031
QY 1082 GTGGTGTAGGATGATTTTTCATTTCTTTCCTTGAAGCTTACTTCAAGTACACA 1141
DB 1032 GGGGTATCAGTGAAGATTTCTTATTTCTTTCCTTGAAGCTTACTTCAAGTACACA 1084
QY 1142 AACTCTTCATCAGGCGCATGATCTATGAGACTCTTATGAGATATCTGGGTGATTTG 1201
DB 1085 ---GCTCTTTTCTGACAAAGTCAAGGCGCTGTCTTCAAGTCTGTCTGATGACAT---- 1137
QY 1202 ACCCAAAACATCTTCCAAAGCATTAATATCCAAATCATGCGTGTATGTTTTAATCAGC 1261
DB 1138 -CACAAGCATTCTCTCAAAAACATTATATCTCAGGACATGCTGTATGTTTCA----- 1191
QY 1262 AGAAGCATGTTTTATGTTGTACAAAAGAAATGTATGAGTGGGAGTGAAGATATAG 1321
DB 1192 -----CTGTCCGCTGTTTTTCACTTGTATGTGAAGGCTTGGGCTGGGATTTGA 1245
QY 1322 ACCATGATGATGACCTTCAAGCTTATTAATTAAGATCTTAAATGGGACAGAGACT 1381
DB 1246 AGAATGACATGATGCTCTGAGGATTTTCAATTAAGATTTTAAATTCAGATGAGACT 1305
QY 1382 GTGAACAAGACACCTTAATATGAGTGTATGATGAGCAATCTTCTGAAAACGCA 1441
DB 1306 ACAGAAATATCATCTCTGAAAATGATGATGAGCTGCAAGAAAGCAATCCCTGAAAACAG 1365
QY 1442 AACTCTTTT 1450
DB 1366 GACTCTTTT 1374

RESULT 6
US-09-764-891-5477
; Sequence 5477, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5477
; LENGTH: 13862
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-5477

Query Match 10.5%; Score 295.2; DB 10; Length 13862;
Best Local Similarity 74.6%; Pred. No. 5,6e-52;
Matches 397; Conservative 0; Mismatches 133; Indels 2; Gaps 2;

QY 1667 AAACAGATGATTAAGCTGCGAACCCTGCGACACCGCTCATGCTTCTAGCTGCTGGAGG 1726
DB 1508 AAACCTTAATTAAGCTGCGCTGGGCTCTGTGGCTCATGCTGTGATCCAGCACTTTGGAGG 1567
QY 1727 TTGAGAGGAGAGATGAGCTTGAACACAGGTTTCAAGCCAGCCCTGGGCAACATTAACAG 1786
DB 1568 CTGAGGAGAGAGATCACTTGAAGTCAAGAGTTTGAGACCACTGGCCCAATGATGCGAA 1627
QY 1787 ATCTGTCTCTCAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1845
DB 1628 ACCCATCTCTACTATAAAATCAAAAGTTAGCCAGCAAGGCGGCGATGTGTGGTGGC 1687
QY 1846 TCAGGCTGTATCCAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGTCAAGA 1905
DB 1688 TCATGCTGTATCCAGCACTTTGGAGGCGGAGGCGGATCACTGTGTCAAGA 1747
QY 1906 GTTTGAGACCAAGCTGCGCAATGCGCAAAACCCCTGTCTACTCAAAATGCAAAATTA 1965
DB 1748 GTTCAAGACCAAGCTGCGCAATGCGCAAAACCCCTGTCTACTCAAAATGCAAAATTA 1807
QY 1966 GCCAGGCGTGTGAGCAAGCACTGTATCCAGCTACTTGGAGGCTGGAGAGAAAT 2025
DB 1808 GCCAGGCTTTATGTAGTGAGGCGCTTATATCCAGCTACTTGGAGGCTGGAGAGAAAT 1867
QY 2026 CGCTTAACCCAGAGAGTGGAGGTTGCAATGAGTGAATCGTGCCTGCACTCCAGCC 2085
DB 1868 CGTTTGAAGCAGAGAGCGGAGTTTGCAACCGAGCTATGTGCCACTGCACTCCAGCC 1927
QY 2086 TGGGCGCAAGACCAAGCTCTGTCTCAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2145
DB 1928 TGGGCGCATTA-AGCGAGACTCTGTCTCAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1986
QY 2146 AGAGAACATATTGGAGAGAGAGATGGGAAAGCTTGCAGGAATTTGTG 2197
DB 1987 AGAAAAAATTAGCCAGGCTGTGTCATGCACTGTGCTCTGACTTGGG 2038

RESULT 7
US-09-764-891-10204
; Sequence 10204, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 10204
; LENGTH: 13862
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-10204

Query Match 10.5%; Score 295.2; DB 10; Length 13862;
Best Local Similarity 74.6%; Pred. No. 5,6e-52;
Matches 397; Conservative 0; Mismatches 133; Indels 2; Gaps 2;

QY 1667 AAACAGATGATTAAGCTGCGAACCCTGCGACACCGCTCATGCTTCTAGCTGCTGGAGG 1726
DB 1508 AAACCTTAATTAAGCTGCGCTGGGCTCTGTGGCTCATGCTGTGATCCAGCACTTTGGAGG 1567
QY 1727 TTGAGAGGAGAGATGAGCTTGAACACAGGTTTCAAGCCAGCCCTGGGCAACATTAACAG 1786
DB 1568 CTGAGGAGAGAGATCACTTGAAGTCAAGAGTTTGAGACCACTGGCCCAATGATGCGAA 1627
QY 1787 ATCTGTCTCTCAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 1845
DB 1628 ACCCATCTCTACTATAAAATCAAAAGTTAGCCAGCAAGGCGGCGATGTGTGGTGGC 1687
QY 1846 TCAGGCTGTATCCAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGTCAAGA 1905
DB 1688 TCATGCTGTATCCAGCACTTTGGAGGCGGAGGCGGATCACTGTGTCAAGA 1747
QY 1906 GTTTGAGACCAAGCTGCGCAATGCGCAAAACCCCTGTCTACTCAAAATGCAAAATTA 1965
DB 1748 GTTCAAGACCAAGCTTGGACCAATGCGCAAAACCCCTGTCTACTTATAAAATGCAAAATTA 1807
QY 1966 GCCAGGCGTGTGAGCAAGCACTGTATCCAGCTACTTGGAGGCTGGAGAGAAAT 2025
DB 1808 GCCAGGCTTTATGTAGTGAGGCGCTTATATCCAGCTACTTGGAGGCTGGAGAGAAAT 1867
QY 2026 CGCTTAACCCAGAGAGTGGAGGTTGCAATGAGTGAATCGTGCCTGCACTCCAGCC 2085
DB 1688 CGTTTGAAGCAGAGAGCGGAGTTTGCAACCGAGCTATGTGCCACTGCACTCCAGCC 1927
QY 2086 TGGGCGCAAGACCAAGCTCTGTCTCAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2145
DB 1928 TGGGCGCATTA-AGCGAAGCTGTCTCAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1986
QY 2146 AGAGAACATATTGGAGAGAGAGATGGGAAAGCTTGCAGGAATTTGTG 2197
DB 1987 AGAAAAAATTAGCCAGGCTGTGTCATGCACTGTGCTCTGACTTGGG 2038

RESULT 8
US-10-205-428-1003
; Sequence 1003, Application US/10205428
; Publication No. US20030108907A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P117C1
; CURRENT APPLICATION NUMBER: US/10/205,428
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11

Db 43446 AGCGGAGGTTGAGAGGACCAAGATCGGCACTGTATCCAGCTGGGTACAA-AGC 43504
QY 2100 AAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGA 2142
43505 AAGACTCCATCTCAAAAAAAGAGAGAGAGAGA 43547

RESULT 13

US-09-954-531-180
; Sequence 180, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:

APPLICANT: Weaver, Zoe
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc

FILE REFERENCE: 689290-77
CURRENT APPLICATION NUMBER: US/09/954,531

PRIOR FILING DATE: 2002-05-02
PRIOR APPLICATION NUMBER: US/60/233,133

PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,009

PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,034

PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,509

PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: US/60/234,567

NUMBER OF SEQ ID NOS: 1392
SOFTWARE: PatentIn version 3.0

SEQ ID NO 180
LENGTH: 65608
TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: misc_feature
LOCATION: (1)...(65608)

OTHER INFORMATION: n=a,l,g or c
US-09-954-531-180

Query Match 10.0%; Score 281.4; DB 9; Length 65608;
Best Local Similarity 81.2%; Pred. No. 1.2e-48;

Matches 376; Conservative 0; Mismatches 81; Indels 6; Gaps 4;

Db 1682 GCTGGAACCGTGGACAGCTCATAGTCTAGCTGCTGGAGGTTGAGAGGAGAT 1741

43089 GCGGGGTGTGTGCTCAGCTGTATCCAGCACTTGAAGTCTGAGGAGAGAT 43148

QY 1742 GCGTTGACACAGGTGTTCAAGGCGCAGCTGGGCAACATTAAGATCCTGCTCAAA 1801

43149 CCTTAGGCCCAATTTCA-GCCAGCTGGGCAATTTAGGAGACGCACTCTACTA 43207

QY 1802 AAAAAAAAAAAAAAGA-AAAGAGAGGCGCGGCGTGTGCTCAAGCTGTAAATC 1859

43208 AAAAAAAAAAAAAAGAAGAAATTTAGCGCGGCGTGTGCTCAAGCTGTAAATC 43267

QY 1860 CCAAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGTGTGAGAGTTGAGACAGC 1919

43268 CCAAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGTGTGAGAGTTGAGACAGC 43325

QY 1920 TGGCCAACTGGAAGAAACCCGCTGTACTCAAAATGCAAAATTAAGCAGGCGTGTAG 1979

43326 TGGCCAACTGGAAGAAACCCGCTGTACTCAAAATTAAGCAGGCGTGTAG 43385

QY 1980 CAGGCACTGTAACTCCAGCTACTTGGAGGCTGAGGAGAGAAATCGCTTGAACCCAG 2039

43386 CAGGCACTGTAACTCCAGCTACTTGGAGGCTGAGGAGAGAAATCGCTTGAACCCAG 43445

QY 2040 AGGTGAGGTTGAGTGAAGTGAATGATGCGCGTTGCACTCAAGCTGGGCAAGAGC 2099

43446 AGGTGAGGTTGAGTGAAGTGAATGATGCGCGTTGCACTCAAGCTGGGCAAGAGC-AGC 43504

QY 2100 AAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGA 2142
Db 43505 AAGACTCCATCTCAAAAAAAGAGAGAGAGAGA 43547

RESULT 14

US-09-764-877-2857
; Sequence 2857, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:

APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877

PRIOR FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 4031
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2857
LENGTH: 11821
TYPE: DNA

ORGANISM: Homo sapiens
US-09-764-877-2857

Query Match 9.9%; Score 280.2; DB 9; Length 11821;
Best Local Similarity 74.4%; Pred. No. 8e-49;

Matches 386; Conservative 0; Mismatches 118; Indels 15; Gaps 2;

QY 1635 AGTATTTTACTATATTAAGAGTTGTGCAAAACAGATGATTAAGTGGCAACCGTG 1694

4599 AGGAACTGCAAACTATGTTCAGCAGCAAGAAATGAACATGAAGCAGGCAATGATG 4658

QY 1695 GCAACGCTCATAGTTCTAGCTGCTGGAGTTGAGAGGAGAGATGCTTGAACAG 1754

4659 GCTATGCGGTATCCAGCACTTTAGAG-ATTAGCAGAGAGATCACTTGAAGCGAG 4717

QY 1755 GTGTTCAAGGCGCAGCTGGGCAACATTAAGATCTGTCTCAAAAAAAGAGAGAG 1814

4718 GAGTTCAAGACAGTCTGGGCAATATAGTAACTTGTCTCAAAAAAAGAGAGAG 4776

QY 1815 AAAAAAAGAGAGAGGCGCGGTGTGCTCAAGCTGTAAATCCAGATCTTGGAG 1874

4777 -----TTGCGAGGCGGTGTGCTCAAGCTGTAAATCCAGATCTTGGAG 4823

QY 1875 GCGGAGCGGCGGATCACTGTGTGAGAGTTGAGACAGCTGGGCAATGAGCA 1934

4824 GCGGAGCGGCGGATCACTGTGTGAGAGTTGAGACAGCTGGGCAATGAGCA 4883

QY 1935 AACCCTGTGTACTCAAAATGCAAAATTAAGCAGGCGTGTAGCAGCACTGTAAATC 1994

4884 AACCCTGTGTACTCAAAATGCAAAATTAAGCAGGCGTGTAGCAGCACTGTAAATC 4943

QY 1995 CAGACTACTTGGAGGCTGAGGCGAGAGATCGCTTGAACCAAGAGTGGAGTTGAG 2054

4944 CAGACTACTTGGAGGCTGAGGCGAGAGATCGCTTGAACCAAGAGTGGAGTTGAG 5003

QY 2055 TAAGCTGATGCTGCGCTGTGCACTCAAGCTGGGCAAGAGCAAGCTGTCTCAG 2114

5004 TAAGCTGATGCTGCGCTGTGCACTCAAGCTGGGCAAGAGCAAGCTGTCTCAG 5063

QY 2115 AAAAAAAG 2153

5064 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5102

RESULT 15

US-10-242-515-2857
; Sequence 2857, Application US/10242515
; Patent No. US20040009488A1
; GENERAL INFORMATION:

APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC005C1

Job time : 971 secs

```

; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,986
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2857
; LENGTH: 11821
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-515-2857

```

```

Query Match          9.9%; Score 280.2; DB 15; Length 11821;
Best Local Similarity 74.4%; Pred. No. 8e-49;
Matches 386; Conservative 0; Mismatches 118; Indels 15; Gaps 2;

```

```

QY 1635 AGGATATTTACTATATTAAGAGAGTGTGACAAACAGATATTAAGCTGGAAACCGTG 1694
DB 4599 AGGAACTGCAAACTATGTTGCGACAGACAGAAAGAAATGAACTGAGCCATGATG 4658
QY 1695 GCACACGCTCATGATGTTCTAGCTGTTGGAGGTTGAGAGAGATGCTTGAACACAG 1754
DB 4659 GCTCATGCGGTATATCCGACACTTAGGA-ATTAGCAGAGAGATCACTTGAAGCCAG 4717
QY 1755 GTGTTCAAGGCGCAGCTGCGCAACATACAGATCTGTCTCAAAAAAAAAAAAAA 1814
DB 4718 GAGTCAAGACGAGCTGCGCAATATAGTAAGACTTGTCTTACAAAAATGAAAAA- 4776
QY 1815 AAAAGAAAGAGAGGCGCGGCGTGTGCTCAAGCTGTATCCAGCACTTTGGAG 1874
DB 4777 -----TTGGCCAGGCGTGTGCTCAAGCTGTATTCAGCACTTTGGAG 4823
QY 1875 GCCGAGCGGCGCGATCACTGTGTCAAGAGTTGAGCCAGCTGGCCCAATGGCA 1934
DB 4824 GCCGAGCGGCGCGATTACTCTGAGTCAAGAGTTCAAGACCAAGCTGGCCCAATGGCA 4883
QY 1935 AACCCGCTGTATCAAAATGCAAAATTTAGCCAGCGTGTAGCAGGCACTGTATC 1994
DB 4884 AACCCGCGCTCTATAAAATACAAAATTTAGCTGGCGGTGTGAGCACTGTATC 4943
QY 1995 CCAAGCTACTGGAGGCTGAGGCAAGAAATCGCTTGAACCAAGAGGTGAGGTTGAG 2054
DB 4944 CCAAGCTACTGGAGGCTGAGGCAAGAAATCACTTGAATCTGGGAGGTGAGGTTGAG 5003
QY 2055 TAAGCTGAGATCGTGGCGTTGCACTCAAGCTGGGCGCAAGAGCAAGCTGTCTCAG 2114
DB 5004 TGAGCCGCGATCCACCATTAACACTACAGCTGGGCAAGAGAGTGAACTGTCTCA 5063
QY 2115 AAAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGACA 2153
DB 5064 AATTAATTAATTAATTAATTAATTAATTAATTAATA 5102

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 06:30:42 ; Search time 7086 Seconds
(without alignment)

11875.759 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818
Sequence: 1 agagaacacatcaatga.....aaaaaaaaaaaaaa 2818

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

- Database :
- 1: em_estba.*
 - 2: em_estba.*
 - 3: em_estba.*
 - 4: em_estba.*
 - 5: em_estba.*
 - 6: em_estba.*
 - 7: em_estba.*
 - 8: em_estba.*
 - 9: em_estba.*
 - 10: em_estba.*
 - 11: em_estba.*
 - 12: em_estba.*
 - 13: em_estba.*
 - 14: em_estba.*
 - 15: em_estba.*
 - 16: em_estba.*
 - 17: em_estba.*
 - 18: em_estba.*
 - 19: em_estba.*
 - 20: em_estba.*
 - 21: em_estba.*
 - 22: em_estba.*
 - 23: em_estba.*
 - 24: em_estba.*
 - 25: em_estba.*
 - 26: em_estba.*
 - 27: em_estba.*
 - 28: em_estba.*
 - 29: em_estba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	970.2	34.4	1201 13	BX402063 BX402063
2	878.8	31.2	1201 9	AL559877 AL559877
3	847	30.1	872 12	BG758510 BG758510
4	817.4	29.0	854 13	BX464578 BX464578

5	809.4	28.7	953 13	BQ065440 BQ065440
6	799.2	28.4	1052 13	BQ055935 BQ055935
7	781.2	27.7	918 13	BX391067 BX391067
8	780.8	27.7	853 13	BX464579 BX464579
9	777.8	27.6	920 13	BX392040 BX392040
10	739.4	26.2	743 12	BG686133 BG686133
11	705.6	25.0	820 12	BG757089 BG757089
12	705.2	25.0	820 12	BG755526 BG755526
13	681	24.2	942 10	BF975096 BF975096
14	667.4	23.7	633 10	BF975166 BF975166
15	663	23.5	664 12	BG754140 BG754140
16	655.4	23.3	653 12	BG757392 BG757392
17	653.2	23.2	1201 9	AL581406 AL581406
18	627.6	22.3	670 12	BG341546 BG341546
19	626	22.2	956 10	BF664355 BF664355
20	615.6	21.8	1201 13	BX402062 BX402062
21	572.6	20.3	604 10	AM978582 AM978582
22	568	20.2	928 12	BG398364 BG398364
23	562	19.9	570 10	AM504807 AM504807
24	551.4	19.6	843 12	BG758815 BG758815
25	546.8	19.4	1034 12	BG755005 BG755005
26	540	19.2	541 10	BF238155 BF238155
27	498	17.7	439 13	BX283642 BX283642
28	493.6	17.5	511 12	BG770756 BG770756
29	490.8	17.4	535 14	CD707143 CD707143
30	471	16.7	517 12	BG170824 BG170824
31	467.8	16.6	528 9	AA954956 AA954956
32	467.6	16.6	939 12	BG341819 BG341819
33	456.4	16.2	889 12	BG686876 BG686876
34	453.8	16.1	1140 10	BF664352 BF664352
35	446.2	15.8	476 10	AM135547 AM135547
36	442	15.7	442 9	AI016902 AI016902
37	441.4	15.7	464 10	AM452648 AM452648
38	441.8	14.6	1292 11	AK080144 AK080144
39	411.4	14.6	757 9	AJ446140 AJ446140
40	409.8	14.5	696 9	AJ453647 AJ453647
41	406.2	14.4	422 10	AM263139 AM263139
42	405.4	14.4	522 12	BG144705 BG144705
43	369.4	13.1	374 10	AM401901 AM401901
44	366.2	13.0	729 9	AJ450296 AJ450296
45	365.6	13.0	688 9	AJ450317 AJ450317

ALIGNMENTS

RESULT 1
LOCUS BX402063
DEFINITION BX402063 Homo sapiens B CELLS (RAMOS CELL LINE) COR 25-NORMALIZED
Homo sapiens cDNA clone CSDBL012YD18 5-PRIME, mRNA sequence.
ACCESSION BX402063
VERSION BX402063.1
KEYWORDS GI:30626645
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLES Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6672.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgl-bin/custer.cgi?seq=CSDBL012DB09P1ac1cluster=6672.r. Contact :
Peng Liang Email: filiang@life.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DL012DB090P1.
Location/Qualifiers

FEATURES

source

```
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL012YD18"
/cell_line="RAMOS CELL LINE"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match 34.4%; Score 970.2; DB 13; Length 1201;
Best Local Similarity 94.1%; Pred. No. 1.9e-116;
Matches 1004; Conservative 30; Mismatches 29; Indels 4; Gaps 3;

```
QY 1 AGAGAACCTCATTATTAAGTGAATTTCTGCGCCGAGACTTGACGAGGAGCAAGA 60
DB 73 AGAGAACCTCATTATTAAGTGAATTTCTGCGCCGAGACTTGACGAGGAGCAAGA 132
QY 61 AGACACTGACACACCACTATGACAGCCTTGATGAAACCGAGAGAAATTTCTTACCA 120
DB 133 AGACACTGACACACCACTATGACAGCCTTGATGAAACCGAGAGAAATTTCTTACCA 192
QY 121 ATTCAAAAATGTCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 180
DB 193 ATTCAAAAATGTCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 252
QY 181 GAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 240
DB 253 GAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 312
QY 241 CTGACACGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 300
DB 313 CTGACACGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 372
QY 301 CTGACACGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 360
DB 373 CTGACACGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 432
QY 361 GAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 420
DB 433 GAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 492
QY 421 CTGACACGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 480
DB 493 CTGACACGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 552
QY 481 GCAATATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 539
DB 553 GCAATATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 612
QY 540 ACCATGAAAGACTTTCAAGCTGAGGAAAGGCTGAGGAAAGGCTGAGGAAAGGCTGAG 599
DB 613 ACCATGAAAGACTTTCAAGCTGAGGAAAGGCTGAGGAAAGGCTGAGGAAAGGCTGAG 672
QY 600 GAGAGCTTGGCGGAGCTTTTGGCGGAGCTTTTGGCGGAGCTTTTGGCGGAGCTTTTGG 659
DB 673 GAGAGCTTGGCGGAGCTTTTGGCGGAGCTTTTGGCGGAGCTTTTGGCGGAGCTTTTGG 732
QY 660 GTAATTGGGACTTGTATGACAACTTCAGAAATGTCAACAAGTAATATCTCTGCT 719
DB 733 GTAATTGGGACTTGTATGACAACTTCAGAAATGTCAACAAGTAATATCTCTGCT 792
QY 720 GAAAGAGTGGATTAAGAAACAGTCTTCAAGTCTTCTGTTTATTTCTTCAACTCTCA 779
DB 793 GAAAGAGTGGATTAAGAAACAGTCTTCAAGTCTTCTGTTTATTTCTTCAACTCTCA 852
```

```
QY 780 CTTCTTAAAGCTTTACAGAAAAATTTATATACGACTTTTAAAGATCTATGCTT 839
DB 853 CTTCTTAAAGCTTTACAGAAAAATTTATATACGACTTTTAAAGATCTATGCTT 912
QY 840 GAAATATGAGAGGAAACACAGGCTCTGCGGAGGACCTGCTGCAATTTGTCAGTTTGA 899
DB 913 GAAATATGAGAGGAAACACAGGCTCTGCGGAGGACCTGCTGCAATTTGTCAGTTTGA 972
QY 900 TGCACATTTGCTCCCTACTGCGGAAATTAACAAGCTGAGGACCTGGAGGACCTTAAAGTG 959
DB 973 TGCACATTTGCTCCCTACTGCGGAAATTAACAAGCTGAGGACCTGGAGGACCTTAAAGTG 1032
QY 960 TCAACGTTTTCATGACTTTTGTGATGATGAGACAGAGAGTATCTTAAAGCAAT 1019
DB 1033 TCAACGTTTTCATGACTTTTGTGATGATGAGACAGAGAGTATCTTAAAGCAAT 1089
QY 1020 GGTGAGAGCATCAATGTTTATATCAATCCTTATATTTGAT 1066
DB 1090 TSSASGSMHAWSTTTTAAAAAACTTATTTTATTTTGT 1136
```

RESULT 2

AL559877

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

Genoscope

Genoscope - Centre National de Sequenage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr; Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 6672.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DG003DA070P1&cluster=6672.r. Contact :

Feng Liang Email: fliang@life.techn.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DG003DA070P1.

Location/Qualifiers

1..1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DG003YB14"

/issue_type="B CELLS (RAMOS CELL LINE)"

/cell_line="RAMOS CELL LINE"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed

with a NotI-oligo (dT) primer. Five prime end enriched,

double-strand cDNA was digested with Not I and EcoRV

the Not I and EcoRV sites of the pCMVSPORT 6 vector.

Library was not normalized."

ORIGIN

Query Match 31.2%; Score 878.8; DB 9; Length 1201;
Best Local Similarity 90.6%; Pred. No. 1.2e-104;
Matches 930; Conservative 24; Mismatches 67; Indels 5; Gaps 2;

```
QY 20 AAGTAGATTTTCTGCGCTGAGACTTGACGAGGAGGACAGACACTCTGACACCACT 79
DB 32 AAGTAGATTTTCTGCGCTGAGACTTGACGAGGAGGACAGACACTCTGACACCACT 91
```

```
QY      80  ATGACAGCCTCTTATGACCGAGAGATTCTTACCAATGCAAAATGCGCTGG 139
DB      92  ATGACAGCCTCTTATGACCGAGAGATTCTTACCAATGCAAAATGCGCTGG 151
QY      140 GCTAAGGGTCGGCGTGAAGCTTACCTGTGCTACGTAAGAGAGCGTGAACAGTGTCTACA 199
DB      152 GCTAAGGGTCGGCGTGAAGCTTACCTGTGCTACGTAAGAGAGCGTGAACAGTGTCTACA 211
QY      200 TCCCTTTTCACTGACCTTTGGTTATCTTCCCAATPAAAGACGCGTCCGACGTGAATTTGCTC 259
DB      212 TCCCTTTTCACTGACCTTTGGTTATCTTCCCAATPAAAGACGCGTCCGACGTGAATTTGCTC 271
QY      260 TTCCCTCCGCTACATCTCCGACCTGGACCTTAAGACCTTGGCGCGTGTCTACCGGCTCACCTGG 319
DB      272 TTCTCTCCGCTACATCTCCGACCTGGACCTTAAGACCTTGGCGCGTGTCTACCGGCTCACCTGG 331
QY      320 TTCACTCTCTGAGACCCCTGCTACGACTGTGCCGACATGTGGCGGACCTTTCTGCGAGAGG 379
DB      332 TTCACTCTCTGAGACCCCTGCTACGACTGTGCCGACATGTGGCGGACCTTTCTGCGAGAGG 391
QY      380 AACCCCAACCTCAGTGTGAGGATCTTACACCGCGCGCTCTACTCTGTGAGGACCGCAAG 439
DB      392 AACCCCAACCTCAGTGTGAGGATCTTACACCGCGCGCTCTACTCTGTGAGGACCGCAAG 451
QY      440 GCTGACCCCGAGGGGCTGCGGGCGCTGCACCGCGCGCGGTGCAAAATGCCATCATGACC 499
DB      452 GCTGACCCCGAGGGGCTGCGGGCGCTGCACCGCGCGCGGTGCAAAATGCCATCATGACC 511
QY      500 TTCAAGATTAATTTTACTGCTGAGAACTATTTTGTGAAACCAATGAAAGAACTTTCAA 559
DB      512 TTCAAGATTAATTTTACTGCTGAGAACTATTTTGTGAAACCAATGAAAGAACTTTCAA 571
QY      560 GCTGAGGAAAGGGTGCATGAAATTCAGTGTCTCTCCAGACAGCTTCCGCGCATCTT 619
DB      572 GCTGAGGAAAGGGTGCATGAAATTCAGTGTCTCTCCAGACAGCTTCCGCGCATCTT 631
QY      620 TTGCCCCCTGTATGAGGTTGATATGACGAGACGATTTGTAATTTGGGACCTTTGATAG 679
DB      632 TTGCCCCCTGTATGAGGTTGATATGACGAGACGATTTGTAATTTGGGACCTTTGATAG 691
QY      680 CAACCTCCGAGAAATGTCACACAGATGAATATCTCTGTGAAAGACAGTGTATAAAC 739
DB      692 CAACCTCCGAGAAATGTCACACAGATGAATATCTCTGTGAAAGACAGTGTATAAAC 751
QY      740 AGCTCTTAAGTCTTCTGTTTATTTTCACTCACTTCTTCACTTCTTGAAGTTTACAGAA 799
DB      752 AGCTCTTAAGTCTTCTGTTTATTTTCACTCACTTCTTCACTTCTTGAAGTTTACAGAA 811
QY      800 AAAATATTAATACGACTTTTAAAGATCTATGCTTGAAGATGAGAGAAACA 859
DB      812 AAAATATTAATACGACTTTTAAAGATCTATGCTTGAAGATGAGAGAAACA 871
QY      860 GGTCTGCGCAGGAGCTGTCTGCAATTTGTGCAAGTTTGAATGCAACATTTGCCCTACTG 919
DB      872 GGTCTGCGCAGGAGCTGTCTGCAATTTGTGCAAGTTTGAATGCAACATTTGCCCTACTG 931
QY      920 GGAATTAACAGAACTGAGAGACCTGGGAGAGATCCTTAAGTGTCAAGTTTCTATGACTT 979
DB      932 GGAATTAACAGAACTGAGAGACCTGGGAGAGATCCTTAAGTGTCAAGTTTCTATGACTT 986
QY      980 TTAGTAGATGAGACGAGAGATGATCTTAAAGAGATGAGAGATGATCTT 1039
DB      987 WGGWGGDTGRBARBARAGWRATCYMAAAACATGGGAGAGTAAATTTTWTAAATCTTT 1046
QY      1040 TTATAT 1045
DB      1047 TTATAT 1052
```

```
RESULT 3
Bg758510
LOCUS      Bg758510      872 bp      mRNA      linear      EST 15-MAY-2001
```

```
DEFINITION 602712721f1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853069 5',
            mRNA sequence.
ACCESSION  Bg758510
VERSION    Bg758510.1 GI:14069163
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 872)
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
JOURNAL    Contact: Robert Strausberg, Ph.D.
            Email: egsab@remail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LHCMI698 row: 1 column: 06
            High quality sequence stop: 836.
FEATURES
            Location/Qualifiers
                1..872
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:4853069"
                 /tissue_type="primary B-cells from tonsils (cell line)"
                 /lab_host="DH10B (phage-resistant)"
                 /clone_lib="NIH MGC 48"
                 /note="Organ: B-cells; Vector: pOT87; Site 1: XhoI;
                 Site 2: EcoRI; cDNA made by oligo-dT priming.
                 Directionally cloned into EcoRI/XhoI sites using the
                 following 5' adaptor: GGCAAGG(G). Size-selected >500bp
                 for average insert size 1.8kb. Library constructed by Ling
                 Hong in the laboratory of Gerald M. Rubin (University of
                 California, Berkeley) using ZAP-cDNA synthesis kit
                 (Stratagene) and Superscript II RT (Life Technologies).
                 Note: this is a NIH_MGC Library."
ORIGIN
Query Match 30.1%; Score 847; DB 12; Length 872;
Best Local Similarity 99.3%; Pred. No. 2e-100; 5; Indels 1; Gaps 1;
Matches 861; Conservative 0; Mismatches 5;
16 ATTGAATGAGATTTTCTGCGCTGAGACTTGAGGAGGAGCAAGACCTCTGACAC 75
2 ATTGAATGAGATTTTCTGCGCTGAGACTTGAGGAGGAGCAAGACCTCTGACAC 61
76 CACTATGACAGCCTCTTGATGACCGGAGAAATTTCTTACCAATTTGAAATGTCGG 135
62 CACTATGACAGCCTCTTGATGACCGGAGAAATTTCTTACCAATTTGAAATGTCGG 121
136 CTGGGCTAAGGGTCGGCGTGAAGACCTACCTGTGCTACGTAAGAGAGCGTGAAG 195
122 CTGGGCTAAGGGTCGGCGTGAAGACCTACCTGTGCTACGTAAGAGAGCGTGAAG 181
196 TACATCCTTTTCACTGACCTTTGGTTATCTTGCATTAAGAAAGCGTGCACGATGAAT 255
182 TACATCCTTTTCACTGACCTTTGGTTATCTTGCATTAAGAAAGCGTGCACGATGAAT 241
256 GCTCTTCTCCGCTACATCTCGACTGGAACCTTAAGCCTTGCGCGCTGTACCGCTCAC 315
242 GCTCTTCTCCGCTACATCTCGACTGGAACCTTAAGCCTTGCGCGCTGTACCGCTCAC 301
316 CTGCTTCACTCTCTGAGAGCCCTGCTACAGACTGTGCCGACATGTGGCCGACCTTTTGG 375
302 CTGCTTCACTCTCTGAGAGCCCTGCTACAGACTGTGCCGACATGTGGCCGACCTTTTGG 361
376 AGGAACCCCAACCTCAGTGTGAGATCTTACACCGCGCGCTCTACTCTGTGAGAGCG 435
```

Db 362 AGGAAACCCAACTCAGTCTGAGATCTTCAACCGGCGCTCTACTTCTGTAGAACCG 421

Oy 436 CAAGCTGAGCCCGAGGGCTGGCGCGCTGCAACCGCGCGGGCTGCAATAGCCATCAT 495

Db 422 CAAGCTGAGCCCGAGGGCTGGCGCGCTGCAACCGCGCGGGCTGCAATAGCCATCAT 481

RESULT 4	854 bp	mRNA	linear	EST 22-MAY-2001
EX464578/C	EX464578	Homo sapiens B CELLS (RAMOS CELL LINE)	Homo sapiens cDNA	
LOCUS	EX464578	clone CS00G003YB14.3-PRIME, mRNA sequence.		
DEFINITION	EX464578			
ACCESSION	EX464578			
VERSION	EX464578.1	GI:31029653		
KEYWORDS	EST.			
SOURCE	Homo sapiens			
ORGANISM	Homo sapiens (human)			

```

FEATURES
  source
    Location/Qualifiers
      1..854
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS060037B14"
        /tissue_type="B CELLS (RAMOS CELL LINE)"
        /cell_line="RAMOS CELL LINE"
        /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"

```

ORIGIN

/note=Vector: pCMWSPORT 6, 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMWSPORT 6 vector. Library was not normalized."

Query Match 29.0%; Score 817.4; DB 13; Length 854;
Best Local Similarity 96.1%; Pred. No. 1.3e-96;
Matches 816; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

QY	929	GAATGACGAGACCTGGAGGATCCTAAAGTGTCAACGTTTTTCATACATTTAGAGT	988
Db	854	GAATGACGAGACCTGGRGATCCTAAAGTGTCAACGTTTTTCATACATTTAGTGG	795
QY	989	ATGAGACGAGAGGTAGTCTCTAAANAGCATGCTGAGAGATCAATGTTTATATCAA	1048
Db	794	ATGAGRGACGAGGTAGTCTCTAARAGCGTGGGAGRGATCAAAAGTTTTTATKTCRA	735
QY	1049	CATCCTTATATTGATTCATTGATTTGATTTCAAGTGGTGTAGATAGATTTTCTATT	1108
Db	734	CATCCTTATATTTATTAATCAATTGAGTTTACAGTGGTGTAGATAGATTTTCTATT	675
QY	1109	CTTTTCCCTGACGTTTACTTTCAAGTATACAAACCTTCCATCAGGCCATGATCTATA	1168
Db	674	CTTTTCCCTTACGTTTAACTTTCAAGTATACAAACCTTCCATCAGGCCATGATCTATA	615
QY	1169	GGACCTCCTTAATGAGAGTACTGGGTATTGTGACCCCAACCATCTCTCCMAAGCATTA	1228
Db	614	GGACCTCCTTAATGAGAGTCTGGGTATTGTGACCCCAACCATCTCTCCMAAGCATTA	555
QY	1229	ATATCCATCATCGCGCTGTATGTTTATATCAGCAAGACATGTTTATGTTTGACAAA	1288
Db	554	ATATCCAGTCACGCGCTGTATGTTTATATCGCAGAGGACATGTTTATGTTTGACAAA	495
QY	1289	AGAAATTTGATATGGGTGGGGATGGAGGTATAGACATGATGGTCACTTCAAGCTACT	1348
Db	494	AGAAATTTGATATGGGTGGGGATGGAGGTATAGACATGATGGTCACTTCAAGCTACT	435
QY	1349	TTAATTAAGATCTTAAATATGGGACAGAGACTGTGAACAAGACACCTTAATATAGGTT	1408
Db	434	TTAATTAAGATCTTAAATATGGGACAGAGACTGTGAACAAGACACCTTAATATAGGTT	375
QY	1409	GATGCTCGAGTAGCAAACTTCTGGAACGCAAACTCTTTAAGAAAGTCCCTAATTTA	1468
Db	374	GATTTCTGAATAGAAATCTTCTGGAACGCAAACTCTTTAAGAAAGTCCCTAATTTA	315
QY	1469	GAACACCCCAAACTTCAATATCATATATTAGCAAAACATTTGAGAGAGTGTCTTGA	1528
Db	314	GAACAAACCCCAAACTTCAATATCATATATTAGCAAAACATTTGAGAGAGTGTCTTGA	255
QY	1529	TGTTGGGAGAGGAAATCTATTTGCTCTGCTGCTCTCTTCACTCAGAGATGCATATC	1588
Db	254	TGTTGGGAGAGGAAATCTATTTGCTCTGCTCTGCTCTCTTCACTCAGAGATGCATATC	195
QY	1589	AGGTCAAGTTTGCTACATTTTGTATGNTGNTAGCTTCTCCCAAAAGTATTTAACTA	1648
Db	194	AGGTCAAGTTTGTCTACATTTTGTATGNTGNTAGCTTCTCCCAAAAGTATTTAACTA	135
QY	1649	TATAAGAGTGTGACAAAACAGATGTATTAAGCTGGAACGCTGGCACACGCTCATAG	1708
Db	134	TATAAGAGTGTGACAAAACAGATGTATTAAGCTGGAACGCTGGGAAACGCTCATAG	75
QY	1709	TTCTAGCTGCTTGGAGGTTGAGAGGGGAGATGGCTTGAACAACAGGTGTTCAAGCCAG	1768
Db	74	TTCTAGCTGCTTGGAGGTTGAGAGGGGAGATGGCTTGAACAACAGGTGTTCAAGCCAG	15
QY	1769	CTTGGGCAA 1777	
Db	14	CTTGGCAAM 6	

RESULT 5

LOCUS	BO065440	953 bp	mRNA	linear	EST 02-APR-2007
DEFINITION	AGENCOURT.6855061 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5929977				
ACCESSION	BO065440				
VERSION	BO065440.1	GI:19894486			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.				
TITLE	Unpublished (1999)				
JOURNAL	Contact: Robert Strausberg, Ph.D.				
COMMENT	Email: cs9ab@remail.nih.gov Tissue Procurement: Lou Staudt cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LCM2108 row: p column: 10 High quality sequence stop: 634. Location/Qualifiers 1. .953 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5929977" /tissue_type="lymphoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 99" /note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."				
FEATURES	source				
ORIGIN	Query Match 28.7%; Score 809.4; DB 13; Length 953; Best Local Similarity 98.5%; Pred. No. 1.3e-95; Matches 828; Conservative 0; Mismatches 11; Indels 2; Gaps 1;				
QY	17 TTGAAGTGAATTTTCTGCGCTGAGACTTCACGAGGACGAAGAAGACACTCTGACACG 76				
DB	1 TTGAAGTGAATTTTCTGCGCTGAGACTTCACGAGGACGAAGAAGACACTCTGACACG 60				
QY	77 ACTATGACAGCCCTCTTGATGATGAACCGGAGGAAGTTCTTACCAATTCAGAAATGTCGC 136				
DB	61 ACTATGACAGCCCTCTTGATGATGAACCGGAGGAAGTTCTTACCAATTCAGAAATGTCGC 120				
QY	137 TGGGCTAAGGTGGCGCGTGAACCTTACTGTGCTACATGAGGAAGGCGCTGACAGTCT 196				
DB	121 TGGGCTAAGGTGGCGCGTGAACCTTACTGTGCTACATGAGGAAGGCGCTGACAGTCT 180				
QY	197 ACATCCCTTTTCACTGAGACTTTGGTTATCTTGACATTAAGAACGGCTGCACTGGAATTG 256				
DB	181 ACATCCCTTTTCACTGAGACTTTGGTTATCTTGACATTAAGAACGGCTGCACTGGAATTG 240				
QY	257 CTCTTCCTCGCTACATCTGGAAGTGGAGCTTAAGCCCTGGCGCTGCTACCGGCTACG 316				
DB	241 CTCTTCCTCGCTACATCTGGAAGTGGAGCTTAAGCCCTGGCGCTGCTACCGGCTACG 300				
QY	317 TGGTTCAACCTGCTGAAGCCCTGCTGACAGCTGTGCGCGCAATGTGACCGGACTTTCTGGGA 376				
DB	301 TGGTTCAACCTGCTGAAGCCCTGCTGACAGCTGTGCGCGCAATGTGACCGGACTTTCTGGGA 360				

QY	377	GGGAACCCCAACCTCAGTCTGAGATCTTCAACCGGCGCCTCTACTTCTGTGAGAACCGC	436
Db	361	GGGAACCCCAACCTCAGTCTGAGATCTTCAACCGGCGCCTCTACTTCTGTGAGAACCGC	420
QY	437	AAAGCTGAGCCCGAGGGGCTGCGGCGGTGCAACGCGCGCGGGGTGCAAAATGACATCATG	496
Db	421	AAAGCTGAGCCCGAGGGGCTGCGGCGGTGCAACGCGCGCGGGGTGCAAAATGACATCATG	480
QY	437	ACCTTCAAGATATATTTTACTGCTGGAAATCTTTGTAGAAAACCATGAAAGAACTTTC	556
Db	481	ACCTTCAAGATATATTTTACTGCTGGAAATCTTTGTAGAAAACCATGAAAGAACTTTC	540
QY	557	AAAGCCTGGAGAGGGCTCATGAAATTCAGTTGCTCTCTCCAGACAGCTTCGCGGCATC	616
Db	541	AAAGCCTGGAGAGGGCTCATGAAATTCAGTTGCTCTCTCCAGACAGCTTCGCGGCATC	600
QY	617	CTTTTGCCCTGTATGAGTTGATGACTTACGAGACGCAATTCGTACTTTGGACCTTGA	676
Db	601	CTTTTGCCCTGTATGAGTTGATGACTTACGAGACGCAATTCGTACTTTGGACCTTGA	660
QY	677	TAGCAACTTCCAGGAATGTCAACACCATGAAATATCTCTGCTGAGACAGTGCATAA	736
Db	661	TAGCAAACTTCAGGAATGTCAACACCATGAAATATCTCTGCTGAGACAGTGCATAA	720
QY	737	AACAGTCTTCAAGCTCTCTCTGTTTTATCTTCAACTCTCACTTCTTGTAGAGTTACA	796
Db	721	AACAGTCTTCAAGCTCTCTCTGTTTTATCTTCAACTCTCACTTCTTGTAGAGTTAC	780
QY	797	GAATAAATATTTATATATAGACTCTTTAAAAA-GATCTATGCTTGAATAATAGAAAGGA	854
Db	781	GAATAAATATTTATATATAGACTCTTTAAAAAAGATCTTGAATAAATAGAAAGG	840
QY	855	A	855
Db	841	A	841
RESULT 6			
LOCUS	B0055935		
DEFINITION	B0055935	1052 bp	mRNA
ACCESSION	AGNCOURT.6796231	NIH_MGC_99	Homo sapiens cDNA clone IMAGE:5808181
VERSION	B0055935		
KEYWORDS	B0055935.1	GI:19815262	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	NIH-MGC	http://mgc.nci.nih.gov/	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-rc@mail.nih.gov Tissue Procurement: Lou Straub cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LNCM2051 row: m column: 14 High quality sequence stop: 665. Location/Qualifiers		
FEATURES	source		
	1..1052		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5808181"		
	/tissue_type="lymphoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		

/note=lib="NIH_MGC_99"
 EcorI, cDNA made by oligo-dt priming. Directionally cloned
 into EcorI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using Zap-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

ORIGIN

Query Match 28.4%; Score 799.2; DB 13; Length 1052;
 Best Local Similarity 99.2%; Pred. No. 2,6e-94;
 Matches 823; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

17 TTGAAGTGAAGATTTTCTGGCTGAGACTTGGAGGAGCAAGAGACACTGTGACACC 76
 1 TTGAAGTGAAGATTTTCTGGCTGAGACTTGGAGGAGCAAGAGACACTGTGACACC 60

77 ACTATGAGACAGCTCTTGATGAAACCGAGAGAAAGTTCTTTACCAATTCAAAATGTCCGC 136
 61 ACTATGAGACAGCTCTTGATGAAACCGAGAGAAAGTTCTTTACCAATTCAAAATGTCCGC 120

137 TGGGCTTAAGGCTGGGCTGAGACCTTCTGTGTACGTAGTAAGAGGCGGTGACAGTCT 196
 121 TGGGCTTAAGGCTGGGCTGAGACCTTCTGTGTACGTAGTAAGAGGCGGTGACAGTCT 180

197 ACATCTTTTCACTGACCTTGGTATCTTCGCAATAAAGAGGCTCCACAGTGAATTG 256
 181 ACATCTTTTCACTGACCTTGGTATCTTCGCAATAAAGAGGCTCCACAGTGAATTG 240

257 CTCTTCTCCGCTACATCTCGACTGAGACCTAGACCTGAGCGGTGATACCGCTCAC 316
 241 CTCTTCTCCGCTACATCTCGACTGAGACCTAGACCTGAGCGGTGATACCGCTCAC 300

317 TGGTCACTCTCTGAGAGCCCTGCTACGACTGTGCCGACATGTGCGCACTTTCTGCGA 376
 301 TGGTCACTCTCTGAGAGCCCTGCTACGACTGTGCCGACATGTGCGCACTTTCTGCGA 360

377 GGGAAACCCCACTCACTGCTGAGAGATCTTCAACGCGCGCTCTTACTTCTGTAGAGACCGC 436
 361 GGGAAACCCCACTCACTGCTGAGAGATCTTCAACGCGCGCTCTTACTTCTGTAGAGACCGC 420

437 AAGCTGAGAGCCGAGGAGGCTGAGAGGCTGACACCGGCGGAGGTGAATATGCAATCAG 496
 421 AAGCTGAGAGCCGAGGAGGCTGAGAGGCTGACACCGGCGGAGGTGAATATGCAATCAG 480

497 ACCTCAAGATTAATTTTACTGCTGAGAAATCTTTGTAGAAAACCATGAAAGAACTTTC 556
 481 ACCTCAAGATTAATTTTACTGCTGAGAAATCTTTGTAGAAAACCATGAAAGAACTTTC 540

557 AAGGCTGGGAGAGGCTGCTGATAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 616
 541 AAGGCTGGGAGAGGCTGCTGATAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600

617 CTTTGGCCCTGATAGAGGTGATGACTTAAGAGAGCAATTCGATCTTTGGAGATTTGA 676
 601 CTTTGGCCCTGATAGAGGTGATGACTTAAGAGAGCAATTCGATCTTTGGAGATTTGA 660

677 TAGCACTTCAAGAGATGCTACACAGATGAATAATCTGCTGAGAGACAGGAGATAAA 736
 661 TAGCACTTCAAGAGATGCTACACAGATGAATAATCTGCTGAGAGACAGGAGATAAA 720

737 AACAGTCC-TTCAAGTCTTCTGTTTATTTTCTTCAACTCTCACTTTCTTGAATTTAC 795
 721 AACAGTCTTCTCAAGCTTCTGTTTATTTTCTTCAACTCTCACTTTCTTGAATTTAC 780

796 AGAAAAAATTAATTAATTAAGACTTTT-AAAAAGATCTATGCTGAGAAA 844
 781 AGAAAAAATTAATTAATTAAGACTTTTAAAAAGATCTATGCTGAGAAA 830

RESULT 7

EX391067

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

LOCATION/Qualifiers

1. 918

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS01012YD18"

/cell_type="B CELLS (RAMOS CELL LINE)"

/cell_line="RAMOS CELL LINE"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 27.7%; Score 781.2; DB 13; Length 918;

Best Local Similarity 95.0%; Pred. No. 6.1e-92;

Matches 872; Conservative 0; Mismatches 38; Indels 8; Gaps 6;

Db	Accession	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	362	CCCTTGAAGTTACTTTCAAGTAA	CAAAAACCTTCCATCAGGCCATGATCTAATGAGCC	421						
QY	1174	TCCTTATGAGATATCTGGGTGATTTG	GAACCCCAACATCTCTCCAAAGCATTAATATC	1233						
Db	422	TCCTTATGAGATATCTGGGTGATTTG	GAACCCCAACATCTCTCCAAAGCATTAATATC	481						
QY	1234	CAATATGCGGCTATGTTTAACTAG	GAAGAAGCATGTTTTAATGTTGTACAAAAGAG	1293						
Db	482	CAATATGCGGCTATGTTTAACTAG	GAAGAAGCATGTTTTAATGTTGTACAAAAGAG	541						
QY	1294	ATTGTTATGGGTGG--GATGAG	GGTATAGACATGATGTCACCTTCAAGTACTTTA	1351						
Db	542	ATTGTTATGGGTGGGTATGGGAG	GGTATAGACATGATGTCACCTTCAAGTACTTTA	601						
QY	1352	ATAAAGATCTTAAATGGGCA	GGAGACTGTGAACAAACACCTTAATATGGTGTAT	1411						
Db	602	ATAAAGATCTTAAATGGGCA	GGAGACTGTGAACAAACACCTTAATATGGTGTAT	661						
QY	1412	GTCTGAATGAGAAATCTTCTG	GAAGCCAACTCTTTTAAAGAACTCCCTAATTTAGA	1471						
Db	662	GTCTGAATGAGAAATCTTCTG	GAAGCCAACTCTTTTAAAGAACTCCCTAATTTAGA	721						
QY	1472	ACACCCCAAACTTCAATATCA	TATATATAGCAAAATTTGG--AAGAAATGCTTGAATG	1530						
Db	722	ACACCCCAAACTTCAATATCA	TATATATAGCAAAATTTGGAAAGAAATGCTTGAATG	781						
QY	1531	TTGGGAGAGGAAAATCTAT	TGCTCTGCTGGTCTCTTCATCTCAGAAATGCAATCAG	1590						
Db	782	TTGGGAGAGGAAAATCTAT	TGCTCTGCTGGTCTCTTCATCTCAGAAATGCAATCAG	839						
QY	1591	GTCAAGTTTGTCTAATTTT	GTATGTGTATGCTTTCTCCCAAGATATATATCTATA	1650						
Db	840	GTCAAGTTTGTCTAATTTT	GTATGTGTATGCTTTCTCCCAAGATATATATCTATA	898						
QY	1651	TAAGAGATTGTGACAA	1668							
Db	899	TAAGAGATTGTGACAA	1668							
RESULT 8	BX464579									
LOCUS	BX464579	853 bp	mRNA	linear	EST 22-MAY-2003					
DEFINITION	BX464579 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA									
ACCESSION	BX464579									
VERSION	BX464579.1	GI:31031641								
KEYWORDS	EST.									
SOURCE	Homo sapiens (human)									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.									
AUTHORS	Li W.B., Gruber C., Jessee J. and Polayes D.									
TITLE	Full-length cDNA libraries and normalization									
JOURNAL	Unpublished (2001)									
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6672.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi:bin/cluster.cgi?seq=CSIDG001ZF100P1&cluster=6672.r. Contact : Peng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSIDG001ZF100P1. Location/Qualifiers 1..853									
FEATURES	source									
		/organism="Homo sapiens"								
		/mol_type="mRNA"								
		/db_xref="taxon:9606"								

```

/clone="CS0DG003YB14"
/risuse type="B" B CELLS (RAMOS CELL LINE)
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

Query Match	27.7%	Score 780.8	DB 13	Length 853
Best Local Similarity	96.7%	Pred. No. 7.2e-92		
Matches	801	Conservative	6	Mismatches 20
			Indels	1
			Gaps	1
QY	20	AAGTGAATTTTTCTGACCTGAGACTTGCAGAGGAGGCAAGAGACACTCTGAGACACACT	79	
Db	27	AAGCAGGCTGGTACCGGTCCGGAATTCGCCGGATGCGAAGMGACACTCTGAGACACACT	86	
QY	80	ATGAGCAGGCTCTTGATGAAACCGAGGAAGTTTCTTACCAATTCAAAATGTCCGCTGG	139	
Db	87	ATGACAGGCTCTTGATGAAACCGGA-GAAGTTCTTTACCAATCAAAAATGTCCGCTGG	145	
QY	140	GCTAAGGGTGGCGGTGAGACCTACCTGCTGCTAGCTGAAAGGCGGTGACAGTGCATCA	199	
Db	146	GCTAAGGGTGGCGGTGAGACCTACCTGCTGCTAGCTGAAAGGCGGTGACAGTGCATCA	205	
QY	200	TCCCTTTCACTGACCTTGGTTATCTTTCGCAATATAGAACGGCTGCCACGTGGAATTGCTC	259	
Db	206	TCCCTTTCACTGACCTTGGTTATCTTTCGCAATATAGAACGGCTGCCACGTGGAATTGCTC	265	
QY	260	TTCCCTCCGCTACATCTTGAGACTGAGACCTAGACCCTGCGCTGCTTACCGCGTCACTGG	319	
Db	266	TTCCCTCCGCTACATCTTGAGACTGAGACCTAGACCCTGCGCGCTGCTACCGCGTCACTGG	325	
QY	320	TTCACTCTCTGAGAGCCCTGCTGTAAGACTGTGACCCTGATGTGCGGCACTTTCTGCGAGGG	379	
Db	326	TTCACTCTCTGAGAGCCCTGCTGTAAGACTGTGACCCTGATGTGCGGCACTTTCTGCGAGGG	385	
QY	380	AAACCCCACTCAATCTGAGAGATTCTTACACCGCGCGCTCTTACTTCTGTGAGAACCGGAG	439	
Db	386	AAACCCCACTCAATCTGAGAGATTCTTACACCGCGCGCTCTTACTTCTGTGAGAACCGGAG	445	
QY	440	GCTGAGCCGAGGGGGCTGCGCGCGCTGCAACGCGCGGGGGTGCATAATGACATCATGACC	499	
Db	446	GCTGAGCCGAGGGGGCTGCGCGCGCTGCAACGCGCGGGGGTGCATAATGACATCATGACC	505	
QY	500	TTCAAAATTTATTTTACTGCTGAGAACTTTGTAGAAAACATGAAAGAACTTTCAA	559	
Db	506	TTCAAAATTTATTTTACTGCTGAGAACTTTGTAGAAAACATGAAAGAACTTTCAA	565	
QY	560	GCTGAGAGGGCTGCATGAAATTCAGTTGCTCTCAGACACGCTTGGCGCATCCTT	619	
Db	566	GCTGAGAGGGCTGCATGAAATTCAGTTGCTCTCAGACACGCTTGGCGCATCCTT	625	
QY	620	TTGGCCCTGTATGAGGTGATGACTTACGAGACGCAATTTGTAATTTTGGCACTTTGATAG	679	
Db	626	TTGGCCCTGTATGAGGTGATGACTTACGAGACGCAATTTGTAATTTTGGCACTTTGATAG	685	
QY	680	CAACTTCAGGAATGTGACACAGATGAATATCTGCTGGAAGACAGTGAATGAATAAC	739	
Db	686	CAACTTCAGGAATGTGACACAGATGAATATCTGCTGGAAGACAGTGAATGAATAAC	745	
QY	740	AGTCCTCAAGTCTTCTGTGTTTATCTTCAACTCTCACTTCTTGAAGTTTACGAA	799	
Db	746	AGTCCTCAAGTCTTCTGTGTTTATCTTCAACTCTCACTTCTTGAAGTTTACGAA	805	
QY	800	AAAAATATTATATACACTCTTTTAAAAAGATCTATGCTTGAATAATAG	847	
Db	806	AAAAATATTATATACACTCTTTTAAAAAGATCTATGCTTGAATAATAG	853	

EX392040/c
 LOCUS EX392040 920 bp mRNA linear EST 13-MAY-2003
 DEFINITION Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 ACCESSION Homo sapiens cDNA clone CS0DL012YD18 3-PRIME, mRNA sequence.
 VERSION EX392040
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 920)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by life technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6672.r for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BA1048ZH08_CS04576_1cluster=6672.r.
 Contact: Peng Liang Email: liang@life.com URL:
 http://fulllength.invitrogen.com/Invitrogen/Corporation 1600
 Faraday Avenue Genoscope sequence ID: CS0BA1048ZH08_CS04576_1.
 Location/Qualifiers
 1..920
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL012YD18"
 /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /cell_line="RAMOS CELL LINE"
 /clone_11b="Homo sapiens B CELLS (RAMOS CELL LINE) COT
 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dt)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 27.6%; Score 777.8; DB 13; Length 920;
 Best Local Similarity 95.2%; Pred. No. 1.7e-91;
 Matches 875; Conservative 0; Mismatches 36; Indels 8; Gaps 7;

369 TTCTGCGAGGGAACCCCACTCAGTCTGAGATCTTACCGCGGCTCTACTTCTGTG 428
 |||||
 911 TTTTGGAGAGGAAACCCCACTCAGTCTGAGG-TCTTCCCGCGGCTCTA-TTTTGTG 854
 |||||
 429 AGGACCGCAAGGCTGAGCCCGAGGCGTGGGGGTGACGCGCGGGGGTGAATAG 488
 |||||
 853 AGGCGCGCAAGG--TGAGCCCGAGGCGTGGCGGNTTGA-CGCGCGGGGGTGAATAG 797
 |||||
 489 CCATCATGACCTTCAAGATATTTTACTGCTGATCTTTGATGAAACCATGAAA 548
 |||||
 796 CCATCATGACC-TCAAGATATTTTACTGCTGG-ATCTTNTGTGAAAAACATG-AA 740
 |||||
 549 GAACTTTCAAGGCTGGGAAAGGGCTGATGAAATTCAGTTGCTCTCAGACAGCTTC 608
 |||||
 739 GACTTTCAAGGCTGGGAAAGGGCTGATGAAATTCAGTTGCTCTCAGACAGCTTC 680
 |||||
 609 GGGGCAATCCTTTGGCCCTGTATGAGGTGAGACCTTAGAGACGATTTCTATTGG 668
 |||||
 679 GGGGCAATCCTTTGGCCCTGTATGAGGTGAGACCTTAGAGACGATTTCTATTGG 620
 |||||
 669 GACTTGATGACCACTTCCAGATGTCAACGATGAAATATCTCTGCTGAAAGAGT 728
 |||||
 619 GACTTGATGACCACTTCCAGATGTCAACGATGAAATATCTCTGCTGAAAGAGT 560
 |||||
 729 GGATAAAAACAGTCTTCAAGTCTTCTGTTTATTCTTCAACTCTCACTTTCTTAG 788
 |||||

Db 559 CGATAAAAACAGTCTTCAAGTCTTCTGTTTATTCTTCAACTCTCACTTTCTAG 500
 Qy 789 AGTTTACAGAAAAATTTATTATACGACTCTTTAAAAAGATCTATCTTGAATAAG 848
 Db 499 AGTTTACAGAAAAATTTATTATACGACTCTTTAAAAAGATCTATCTTGAATAAG 440
 Qy 849 GAGGAAACAGGCTCGGCGAGGAGGCTGTGAATTGGTGAAGTTGAATGCAATT 908
 Db 439 GAAAGAAACAGGCTCGGCGAGGAGGCTGTGAATTGGTGAATGCAATT 380
 Qy 909 GTCCCTACTGAGAAATACAGAACTGAGACCTGGAGACATCTTAAAGTCAAGTTT 968
 Db 379 GTCCCTACTGAGAAATACAGAACTGAGACCTGGAGACATCTTAAAGTCAAGTTT 320
 Qy 969 TTCTATGACTTTAGTAGAGATGAGACGAGAGATGATCTTAAAAAGATGGTAGAG 1028
 Db 319 TTCTATGACTTTAGTAGAGATGAGACGAGAGATGATCTTAAAAAGATGGTAGAG 260
 Qy 1029 ATCAATGTTTTTATATCAACATCTTTATTTATTTATTTATTTATTTATTTATTT 1088
 Db 259 ATCAATGTTTTTATATCAACATCTTTATTTATTTATTTATTTATTTATTTATTT 200
 Qy 1089 TAGGATGAGATTTTCTATTTCTTTCCCTGACGTTTACTTCAAGTACACAACTCTT 1148
 Db 199 TAGGATGAGATTTTCTATTTCTTTCCCTGACGTTTACTTCAAGTACACAACTCTT 140
 Qy 1149 CCATCAGGCGCATGATCTATAGAGCCCTCTATAGAGATGATGGGATGGAGCCCAA 1208
 Db 139 CCATCAGGCGCATGATCTATAGAGCCCTCTATAGAGATGATGGGATGGAGCCCAA 80
 Qy 1209 ACCATCTTCCAAAGATTAATATCCATATGCGCTGATGTTTAAATGAGCAAGA 1268
 Db 79 ACCATCTTCCAAAGATTAATATCCATATGCGCTGATGTTTAAATGAGCAAGA 20
 Qy 1269 TGTTTTATGTTGTACAA 1287
 Db 19 GTTTTATGTTGTACAAAAA 1

RESULT 10
 LOCUS BG686133 743 bp mRNA linear EST 01-MAY-2001
 DEFINITION 602638412F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4766234 5',
 mRNA sequence.
 ACCESSION BG686133
 VERSION BG686133.1 GI:13917530
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: ggaabs-remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.llnl.gov
 Plate: L10K1626 row: g column: 03
 High quality sequence stop: 740.
 Location/Qualifiers
 1..743
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4766234"
 /tissue_type="primary B-cells from tonsils (cell line)"

/lab host="DH10B (phage-resistant)"
 /clone lib="NIH_MGC_48"
 /note=TOrgan: B-cells; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGG(6). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match	26.2%	Score 739.4	DB 12	Length 743
Best Local Similarity	99.9%	Pred. No. 1.8e-86		
Matches 740	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	4	GAACATCATTAATTAAGTGAATTTCTGACCTGAGACTTGACAGGAGGCAAGA	63	
DB	2	GAACATCATTAATTAAGTGAATTTCTGACCTGAGACTTGACAGGAGGCAAGA	61	
QY	64	CACCTGGAACACCATGATGAGACGCTCTGATGAACCGAGAGATTCTTACCAAT	123	
DB	62	CACCTGGAACACCATGATGAGACGCTCTGATGAACCGAGAGATTCTTACCAAT	121	
QY	124	CAAAAATGTCCTGGGCTTAAGGCTGGGCTGAGACTTACTGCTACGTAAGTGAAG	183	
DB	122	CAAAAATGTCCTGGGCTTAAGGCTGGGCTGAGACTTACTGCTACGTAAGTGAAG	181	
QY	184	GGGTGACAGTGCATCATCTTTTCACTGACCTTGGTATCTTCCCAATTAAGAAGGCTG	243	
DB	182	GGGTGACAGTGCATCATCTTTTCACTGACCTTGGTATCTTCCCAATTAAGAAGGCTG	241	
QY	244	CCACGTGAATGCTCTTCTCCGCTACATCTCGACTGAGACCTAGACCCCTGCGCTG	303	
DB	242	CCACGTGAATGCTCTTCTCCGCTACATCTCGACTGAGACCTAGACCCCTGCGCTG	301	
QY	304	CTACCGCGTACCTGCTTCACTCTGAGACCCCTGCTACGACTGTCGCCGACATGTGGC	363	
DB	302	CTACCGCGTACCTGCTTCACTCTGAGACCCCTGCTACGACTGTCGCCGACATGTGGC	361	
QY	364	CGACTTCTGCGAGGAGAACCCCACTCAGTCTGAGATCTTACCGCGGCTCTTACTT	423	
DB	362	CGACTTCTGCGAGGAGAACCCCACTCAGTCTGAGATCTTACCGCGGCTCTTACTT	421	
QY	424	CTGTGAGACCGGAGGCTGAGCCGAGGCGCTGCGGCGCTGACCGCGCGGCGTGC	483	
DB	422	CTGTGAGACCGGAGGCTGAGCCGAGGCGCTGCGGCGCTGACCGCGCGGCGTGC	481	
QY	484	AATAGCATCATGACCTTCAAGATTAATTTTACTGCTGAATATCTTTGTGAAGAAC	543	
DB	482	AATAGCATCATGACCTTCAAGATTAATTTTACTGCTGAATATCTTTGTGAAGAAC	541	
QY	544	TGAAGAACTTTCAAGGCTGAGAGGCTGCTGAATAAATTCAGTTCCTCTCCAGCA	603	
DB	542	TGAAGAACTTTCAAGGCTGAGAGGCTGCTGAATAAATTCAGTTCCTCTCCAGCA	601	
QY	604	GCTTGGGCGATCTTTTGGCCCTGATGAGGTTGATGACTTACGAGCGCATTTCCGAC	663	
DB	602	GCTTGGGCGATCTTTTGGCCCTGATGAGGTTGATGACTTACGAGCGCATTTCCGAC	661	
QY	664	TTTGGGACTTTGATAGCAATCTCAGAGATGTCACACGATGAATAATCTCTGCTGAG	723	
DB	662	TTTGGGACTTTGATAGCAATCTCAGAGATGTCACACGATGAATAATCTCTGCTGAG	721	
QY	724	ACAGTGAATAAAAACAATGTC 744		
DB	722	ACAGTGAATAAAAACAATGTC 742		

RESULT 11
 BG757089
 LOCUS

BG757089 820 bp mRNA linear EST 15-MAY-2001

DEFINITION	602715124F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:485517 5', mRNA sequence.
ACCESSION	BG757089
VERSION	BG757089.1
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 820) NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
AUTHORS	Contact: Robert Straubeberg, Ph.D. Email: cgabds-remail.nih.gov Tissue Procurement: Louis M. Staudt, M.D., Ph.D. cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LNCM704 row: 0 column: 06 High quality sequence stop: 675.
FEATURES	Location/Qualifiers 1..820 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:485517" /tissue_type="primary B-cells from tonsils (cell line)" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_48" /note="TOrgan: B-cells; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(6). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN	
Query Match	25.0%
Best Local Similarity	96.5%
Matches 753	Conservative 0
	Mismatches 24
	Indels 3
	Gaps 3
QY	1
DB	6
QY	61
DB	66
QY	121
DB	126
QY	181
DB	186
QY	241
DB	246
QY	301
DB	306
QY	361

Db 366 GCGCGACTTTCTGGAGGAGAACCCCAACTCTGAGGAGTCTTCAACCGCGGCTCTA 425
QY 421 CTTCTGTAGAGACCGGAGGCTGAGCCCGAGGGGCTGCGGCTGCAACCGCGGGGT 480
Db 426 CTTCTGTAGAGACCGGAGGCTGAGCCCGAGGGGCTGCGGCTGCAACCGCGGGGT 485
QY 481 GCAATATGCATATGACCTTCAAAATATATTTTCTGCTGAATACCTTTGTGAAAA 540
Db 486 GCAATATGCATATGACCTTCAAAATATATTTTCTGCTGAATACCTTTGTGAAAA 545
QY 541 CCAATGAAGAACTTCAAGCTGGAAGGGCTGCATGAAATTCAGTCTGCTCCAG 600
Db 546 CCAATGAAGAACTTCAAGCTGGAAGGGCTGCATGAAATTCAGTCTGCTCCAG 605
QY 601 ACAAGCTTGGCGGATCTTTTGGCCCTGTATGAGGTTGATGACTTACGAGACGATTTGG 660
Db 606 ACAAGCTTGGCGGATCTTTTGGCCCTGTATGAGGTTGATGACTTACGAGACGCA-TTGG 664
QY 661 TACTTTGGGACTTTGATAGCACTTCCAGAAATGTCAACA-CGATGAATATCTCTGCT 719
Db 665 TACTTTGGGACTTTGATAGCACTTCCAGAAATGTCAACA-CGATGAATATCTCTGCT 724
QY 720 GAAGACAGTGGATTA-AAAAAGTCTCAAGTCTTCTGTTTATTTCTCAACTGTC 778
Db 725 GAAGACAGTGGATTA-AAAAAGTCTCAAGTCTTCTGTTTATTTCTCAACTGTC 784

RESULT 12
Bg755526 820 bp mRNA linear EST 15-MAY-2001
LOCUS 602716206F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4856568 5',
DEFINITION mRNA sequence.
ACCESSION Bg755526
VERSION Bg755526.1 GI:14066179
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 820)
REFERENCE NIH-MGC http://mgs.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LINC1707 row: k column: 01
High quality sequence stop: 638.
Location/Qualifiers
1. 820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4856568"
/issue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 48"
/note="Organ: B-cells; Vector: pOT8; Site 1: XhoI;
Site 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN

Query Match 25.0%; Score 705.2; DB 12; Length 820;
Best Local Similarity 94.5%; Pred. No. 4.3e-82;
Matches 775; Conservative 0; Mismatches 38; Indels 7; Gaps 4;

QY 634 GGTGATGACTTACGAGAGCATTTGCTGATTTGGGCTTGTATGCACTTCAGGAAT 693
Db 2 GGTGATGACTTACGAGAGCATTTGCTGATTTGGGCTTGTATGCACTTCAGGAAT 60
QY 694 GTCAACACGATGAAATATCTGCTGAGACAGTGTATTAATAAGATCTTCAAGTCT 753
Db 61 GTCAACACGATGAAATATCTGCTGAGACAGTGTATTAATAAGATCTTCAAGTCT 120
QY 754 TCTCTGTTTATTTCTTCAACTCTCACTTCTTGAAGTTTACGAAAAATTTATATA 813
Db 121 TCTCTGTTTATTTCTTCAACTCTCACTTCTTGAAGTTTACGAAAAATTTATATA 180
QY 814 CGACTCTTTAAAAAGATCTATGCTTCTGAAATAGAGAGGACACAGTCTGGCCGAGGA 873
Db 181 CGACTCTTTAAAAAGATCTATGCTTCTGAAATAGAGAGGACACAGTCTGGCCGAGGA 240
QY 874 CGTCTGCAATTTGGTGCAGTTTGAATGCAACATTTGCCCTACTGGGAAATACAGACT 933
Db 241 CGTCTGCAATTTGGTGCAGTTTGAATGCAACATTTGCCCTACTGGGAAATACAGACT 300
QY 934 GCAGACCTGGAGCATCTTAAAGTCAACGTTTCTATGACTTTAGTGAATGAG 993
Db 301 GCAGACCTGGAGCATCTTAAAGTCAACGTTTCTATGACTTTAGTGAATGAG 360
QY 994 AGCAGAGGTGATCTTAAAAAGATGTGTAGAGGATCAATGTTTATATCAACATCC 1053
Db 361 AGCAGAGGTGATCTTAAAAAGATGTGTAGAGGATCAATGTTTATATCAACATCC 420
QY 1054 TTATATTTGATGATTTAGTTTAAAGTGTAGATGATGATTTTCTATCTTT 1113
Db 421 TTATATTTGATGATTTAGTTTAAAGTGTAGATGATGATTTTCTATCTTT 480
QY 1114 CCTTGAAGTTTACTTCAAGTAAACAACAATCTTCCATGAGCCATGATCTATAGACC 1173
Db 481 CCTTGAAGTTTACTTCAAGTAAACAACAATCTTCCATGAGCCATGATCTATAGACC 540
QY 1174 TCTTAATGAGAGTATCTGGGTGATGTGACCCCAACATCTCTCCAAAGCATTAATATC 1233
Db 541 TCTTAATGAGAGTATCTGGGTGATGTGACCCCAACATCTCTCCAAAGCATTAATATC 600
QY 1234 CAATCATGCGCTGATGTTTATATCAGCAGACAGATGTTTATGTTGTACAAAGAG 1293
Db 601 CAATCATGCGCTGATGTTTATATCAGCAGACAGATGTTTCCATGTCCGTACAAAGAG 660
QY 1294 ATTGTATGGGTGGGAGATGAGGAT---AGACATGATGATGATCACTTCAAG--CTACT 1348
Db 661 ATTGTATGGGTGGGAGATGAGGATGCGCGTCTCAGACATGATGATGATCACTTCAAGGCTA 720
QY 1349 TTAATTAAGATCTTAATTAATGGGAGAGGACTGTGAACAAGACACCTTAATAGGTT 1408
Db 721 TTAATTAAGATCTCAATGAGGAGCAAGAGCTGTGAACAAGAAATCTTAATTAAGGTT 780
QY 1409 GATGTCGAGG--TGCAAAATCTTTGGAACGCAAACTCT 1447
Db 781 GATGTCGAGGATTAAGCAAAATCTCTGGAACGCAAACTCT 820

RESULT 13
Bg975096 693 bp mRNA linear EST 22-JAN-2001
LOCUS 602245679F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4336722 5',
DEFINITION mRNA sequence.
ACCESSION Bg975096
VERSION Bg975096.1 GI:12342311
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteleia; Primates; Cetartihni; Homnidae; Homo.
1 (bases 1 to 693)
NTH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

FEATURES	Location/Qualifiers
source	1. .693

ORIGIN

Query Match	24.2%	Score 681,	DB 10,	Length 693,
Best Local Similarity	99.9%	Pred. No. 6.5e-79;		
Matches 692;	Conservative	0;	Mismatches 0;	Indels 1; Gaps 1

Qy	30	CTACCGGATCACTGGATTACCTCCCTGGAGACCCCTGTACAGCATGTGGCCGACATGTAGG	363
Db	1	CTACCGGATCACTGGATTACCTCCCTGGAGACCCCTGTACAGCATGTGGCCGACATGTAGG	60
Qy	364	CGACTTTCTGCGAGGGAAACCCCACTCAGTCTGAGATCTTCAACGCGCGCTTACTT	423
Db	61	CGACTTTCTGCGAGGGAAACCCCACTCAGTCTGAGATCTTCAACGCGCGCTTACTT	120
Qy	424	CTATGAGGACCGGAAGGCTGACCCGAGGGGCTGGGGGCTGTGACCGCGCCGGGGTGA	483
Db	121	CTATGAGGACCGGAAGGCTGACCCGAGGGGCTGGGGGCTGTGACCGCGCCGGGGTGA	180
Qy	484	AATAGCATCATGACCTTCMAAGATTATTTTACTGCTGGAATCTTTTGTAGAAAACA	543
Db	181	AATAGCATCATGACCTTCMAAGATTATTTTACTGCTGGAATCTTTTGTAGAAAACA	240
Qy	544	TGAAAGATCTTTCAAAAGCCTGGGAAGGCGTCATGAAATTCAGTTGCTCTTCAGACA	603
Db	241	TGAAAGATCTTTCAAAAGCCTGGGAAGGCGTCATGAAATTCAGTTGCTCTTCAGACA	300
Qy	604	GCTTGGCGCATCTTTTGCCTCCCTGATAGGTTGATGACTTACGAGACGCAATTGGTAC	663
Db	301	GCTTGGCGCATCTTTTGCCTCCCTGATAGGTTGATGACTTACGAGACGCAATTGGTAC	360
Qy	664	TTTGGACCTTGTATGACAACTTCCAGGAATGTACACACGATGAATATCTTGTAGAG	723
Db	361	TTTGGACCTTGTATGACAACTTCCAGGAATGTACACACGATGAATATCTTGTAGAG	420
Qy	724	ACAGTGGATTAATAAACAATCTTCAAGCTTCTGTTTTTATCTTCAACTCTCACTTT	783
Db	421	ACAGTGGATTAATAAACAATCTTCAAGCTTCTGTTTTTATCTTCAACTCTCACTTT	480

Qy	CTAGAGCTTACAGAAAAAATTTATATACAGCTTTAAAAAGATCTATGCTGAAA	843
Db	CTTAGGCTTACAGAAAAAATTTATATACAGCTTTAAAAAGATCTATGCTTGGAAA	540
Qy	ATAGAGAGAGAAACAGAGCTTGCGCCAGGAGCGTCTGCAATTGGTCAGTTTGAATGCA	903
Db	ATAGAGAGAGAAACAGAGCTTGCGCCAGGAGCGTCTGCAATTGGTGCAG-TTTGAATGCA	599
Qy	ACATTTGTCCCCCTACTGCGAATTACAGAACTGCAGAGCCTGGAGCATCTTAAAGTGCAA	963
Db	ACATTTGTCCCCCTACTGCGAATTACAGAACTGCAGAGCCTGGAGCATCTTAAAGTGTCAA	659
Qy	CGTTTTCATGACTTTTAGGAGGATAGAGC	996
Db	CGTTTTCATGACTTTTAGGAGGATAGAGC	692

RESULT	14
LOCUS	Bf975166
DEFINITION	Bf975166 942 bp mRNA linear EST 22-JAN-2001 60224657.F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4335639 5', mRNA sequence.
ACCESSION	Bf975166
VERSION	Bf975166
KEYWORDS	Bf975166.1 GI:12342381
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens

FEATURES
source

ORIGIN

Query Match	23.7%	Score 667.4;	DB 10;	Length 942;
Best Local Similarity	91.3%	Pred. No. 3e-77;		
Matches 799; Conservative	0;	Mismatches 61;	Indels 15;	Gaps 8;

Oy	4	GAACCAATCATTTATGGAATGAGATTTTCTGGCCCTGGAGACTTGGACGGGAAGCAAGAAGA	63
Dd	2	GAACCATCATTTATGGAATGAGATTTTCTGGCCCTGGAGACTTGGACGGGAAGCAAGAAGA	61
Oy	64	CACCTTGGACACCATGATGACAGCCTCTTGATGAAACCGGAGGAAGTTTCTTTTCCAA-T	122

```

Db      62  CACTGAGACACCATATGAGACGCTCTGATGAACCGAGAGATTCTTACCAAGT 121
Qy      123  TCAAAATGTCCTGCTGGGCTAAGGCTGGGCTGAGACTTACTTGTCTACGATGAGAA 182
Db      122  TCAAAATGTCCTGCTGGGCTAAGGCTGGGCTGAGACTTACTTGTCTACGATGAGAA 181
Qy      183  GGGGTGACAGTGTACATCTTCTCACTGAGACTTGTCTATCTGCAATAGAAAGGCT 242
Db      182  GGGGTGACAGTGTACATCTTCTCACTGAGACTTGTCTATCTGCAATAGAAAGGCT 241
Qy      243  GCCACGTGAGATGCTCTCTCCGCTACATCTGAGACTGAGACTGAGACTGAGACT 302
Db      242  GCCACGTGAGATGCTCTCTCCGCTACATCTGAGACTGAGACTGAGACTGAGACT 301
Qy      303  GCTACCGGCTGACCTGCTGCTGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 362
Db      302  GCTACCGGCTGACCTGCTGCTGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 361
Qy      363  CGGACTTCTGCGAGGAGAACCCGACCTGAGCTGAGATCTTCAACCGGCGCTCTACT 422
Db      362  CGGACTTCTGCGAGGAGAACCCGACCTGAGCTGAGATCTTCAACCGGCGCTCTACT 421
Qy      423  TCTGTGAGACCGCAAGGCTGAGAGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 482
Db      422  TCTGTGAGACCGCAAGGCTGAGAGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 481
Qy      483  AAATAGCATCATGACCTTCAAGATTTATTTACTGCTGGAATCTTTTGTAGAAAC 542
Db      482  AAATAGCATCATGACCTTCAAGATTTATTTACTGCTGGAATCTTTTGTAGAAAC 541
Qy      543  ATGAAAGACTTTCAAGCTGAGAGGCTGCTGATGAAATTCAGTTGCTCTCTCAAGC 602
Db      542  ATGAAAGACTTTCAAGCTGAGAGGCTGCTGATGAAATTCAGTTGCTCTCTCAAGC 601
Qy      603  AGCTGCGGCGCATCTTTTGGCCCTGTATGA--GGTATGACTTTACGAGAGCATTTG 661
Db      602  AGCTGCGGCGCATCTTTTGGCCCTGTATGAAGGCTGTATGACTTTACGAGAGCATTTG 661
Qy      662  ACTTTGGGACTTTATGAGCACTTCCAGAGATGTCACACAGATGAAATATCTGCTGA 721
Db      662  ACTT--GGGACTTTATGAGCA--TTCAGAGATGTCACAC--GATGAAATATCTGCTGA 717
Qy      722  AGACAGTGATTAATAAAACAGTCTTCAAGTCTTCTGTTTATTTCTCAACTCTCACT 781
Db      718  AGACAGTGATTAATAAAAGGCTTAACTACACGAGAGAAATCTT---AAATTTACG 773
Qy      782  TTTCTAGAGTTTCAAAAAAATTTATATACGACTTTTAAAAAGATCTATGCTTGA 841
Db      774  TTTCTTAAAGTTAAGAAAA--ATTTGATTTAGATCTTTTAAAGATTTATGTTTGA 830
Qy      842  AAATAGAGAGAGAGAGAGAGGCTGCGCAGGAGCT 876
Db      831  AA--TATGAGAGAGAGAGGCTTGTGAGAGAGAGT 863

```

```

RESULT 15
Bg754140      664 bp  mRNA  1linear  EST 15-MAY-2001
LOCUS        60279968.F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:484605 5'
DEFINITION   mRNA sequence.
ACCESSION   Bg754140
VERSION      Bg754140.1 GI:14064793
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 664)
AUTHORS     NIH-MGC http://mgs.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.

```

Email: cgaabs-rc@mail.nih.gov
 Tissue Procurement: Louise M. Staedt, M.D., Ph.D.
 CDNA Library Preparation: Ling Hong/Rubin laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LCM1686 row: b column: 22
 High quality sequence stop: 662.
 Location/Qualifiers

FEATURES

```

source
1..664
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:484605"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 48"
/notes="Organ: B-cells; Vector: pOT8; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

```

ORIGIN

```

Query Match      23.5%; Score 663; DB 12; Length 664;
Best Local Similarity 100.0%; Pred. No. 14e-76;
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      634  GGTGATGACTTACGAGACGCTTGTCTGTTGAGACTTTGATGACCACTTCCAGAAAT 693
Db      2  GGTGATGACTTACGAGACGCTTGTCTGTTGAGACTTTGATGACCACTTCCAGAAAT 61
Qy      694  GTACACACAGATGAATATCTCTGCTGAGAGACGTGATTAATAAACAATCTTCAAGTCT 753
Db      62  GTACACACAGATGAATATCTCTGCTGAGAGACGTGATTAATAAACAATCTTCAAGTCT 121
Qy      754  TCTGTGTTTATTTCTTCAACTCTCACTTTTCTTATGAGTTTACGAAATATTTATATA 813
Db      122  TCTGTGTTTATTTCTTCAACTCTCACTTTTCTTATGAGTTTACGAAATATTTATATA 181
Qy      814  CGACTCTTAAAGATCTATGCTTGAATAATGAGAGAGACAGAGTCTGGCCAGGGA 873
Db      182  CGACTCTTAAAGATCTATGCTTGAATAATGAGAGAGACAGAGTCTGGCCAGGGA 241
Qy      874  CGTGTGCAATTTGGTGTGAGTTTGAATGCAACATTTCCCTTACTGAGATTAACAGAACT 933
Db      242  CGTGTGCAATTTGGTGTGAGTTTGAATGCAACATTTCCCTTACTGAGATTAACAGAACT 301
Qy      934  GCAGACCTGGAGACATCTTAAAGTCAACGTTTCTTATGACTTTTATGATGATGAG 993
Db      302  GCAGACCTGGAGACATCTTAAAGTCAACGTTTCTTATGACTTTTATGATGATGAG 361
Qy      994  AGCAGAGTGATGATCCCTAAAGAGCTGAGAGATCAAAATCTTTTATATACATCC 1053
Db      362  AGCAGAGTGATGATCCCTAAAGAGCTGAGAGATCAAAATCTTTTATATACATCC 421
Qy      1054  TTTATTTATGATTCATTTAGTTTAAACAGTGTGATGATGATTTTCTATCTTTT 1113
Db      422  TTTATTTATGATTCATTTAGTTTAAACAGTGTGATGATGATTTTCTATCTTTT 481
Qy      1114  CCTTGAAGTTTACTTTCAAGTAAACAAACCTTCCATCAGGCGCATGATCTATAGACC 1173
Db      482  CCTTGAAGTTTACTTTCAAGTAAACAAACCTTCCATCAGGCGCATGATCTATAGACC 541
Qy      1174  TCTTAAGAGATGATCTGGGTGATTTGAGACCCCAACATCTTCCAAAGCATTAATATC 1233
Db      542  TCTTAAGAGATGATCTGGGTGATTTGAGACCCCAACATCTTCCAAAGCATTAATATC 601

```

Mon Mar 8 10:28:45 2004

us-09-966-880a-7.rst

Page 13

QY 1234 CAAATGCGGCTGTATGTTTAAATCAGCAGAGCATGTTTATGTTGTACAAAGAG 1293
Db 602 CAAATGCGGCTGTATGTTTAAATCAGCAGAGCATGTTTATGTTGTACAAAGAG 661

QY 1294 ATT 1296
Db 662 ATT 664

Search completed: March 7, 2004, 11:56:02
Job time : 7092 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 11:59:52 ; Search time 10902 Seconds

(without alignments)
11203.510 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818
Sequence: 1 agagaccatcatcatga.....aaaaaaaaaaaaaaaaaaaaa 2818Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 2167151695 residues

Word size : 0

Total number of hits satisfying chosen parameters: 692750

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Listing first 45 summaries

Database :

GenBank:
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rdt:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcg_hum:*
40: em_hcg_mus:*
41: em_hcg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	0.7	20	6	AR064875
C 2	20	0.7	20	6	AR080000
C 3	20	0.7	20	6	AR085926
C 4	20	0.7	20	6	AR087520
C 5	20	0.7	20	6	AR093312
C 6	20	0.7	20	6	AR118970
C 7	20	0.7	20	6	AR121692
C 8	20	0.7	20	6	AR123335
C 9	20	0.7	20	6	AR141070
C 10	20	0.7	20	6	AR154115
C 11	20	0.7	20	6	AR164658
C 12	20	0.7	20	6	E12676
C 13	20	0.7	20	6	E12676
C 14	20	0.7	20	6	AR13738
C 15	20	0.7	20	6	AR222466
C 16	20	0.7	20	6	AR236083
C 17	20	0.7	20	6	AR266075
C 18	20	0.7	20	6	AR274394
C 19	20	0.7	20	6	AR305124
C 20	20	0.7	20	6	AR309228
C 21	20	0.7	20	6	AR337145
C 22	20	0.7	20	6	AR343047
C 23	20	0.7	20	6	AR344936
C 24	20	0.7	20	6	AR365970
C 25	20	0.7	20	6	AR382312
C 26	20	0.7	20	6	AR429653
C 27	20	0.7	20	6	AR404876
C 28	20	0.7	20	6	AR4045779
C 29	20	0.7	20	6	AR4045787
C 30	20	0.7	20	6	AR4045790
C 31	20	0.7	20	6	AR104034
C 32	20	0.7	20	6	AR104364
C 33	20	0.7	20	6	AR104368
C 34	20	0.7	20	6	AR115919
C 35	20	0.7	20	6	AR133853
C 36	20	0.7	20	6	AR196224
C 37	20	0.7	20	6	AR196239
C 38	20	0.7	20	6	AR354974
C 39	20	0.7	20	6	AR355810
C 40	20	0.7	20	6	AR355811
C 41	20	0.7	20	6	AR440125
C 42	20	0.7	20	6	AR440140
C 43	20	0.7	20	6	AR465311
C 44	20	0.7	20	6	AR465326
C 45	20	0.7	20	6	AR547087

ALIGNMENTS

RESULT 1
AR064875/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

AR064875
Sequence 5 from patent US 5849480.
AR064875.1 GI:5995091

Unknown.
Unclassified.
1 (bases 1 to 20)
Cros,P., Kurfurst,R., Battall,N. and Pigs,N.
Process and device for assaying a hapten
Patent: US 5849480-A 5 15-DEC-1998;
Location/Qualifiers

20 bp
DNA
linear
PAT 29-SBP-1999

```
source 1.20
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 2
LOCUS AR080000 20 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 83 from patent US 5968524.
ACCESSION AR080000
VERSION AR080000.1 GI:10006735
KEYWORDS
SOURCE Unknown.
ORGANISM Unassigned.
REFERENCE 1 (bases 1 to 20)
AUTHORS Watson,J.D. and Tan,P.L.J.
TITLE Methods and compounds for the treatment of immunologically-mediated
psoriasis
JOURNAL Patent: US 5968524-A 83 19-OCT-1999;
FEATURES
source 1.20
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 3
LOCUS AR085926 20 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 83 from patent US 5985287.
ACCESSION AR085926
VERSION AR085926.1 GI:10012692
KEYWORDS
SOURCE Unknown.
ORGANISM Unassigned.
REFERENCE 1 (bases 1 to 20)
AUTHORS Tan,P., Skinner,M. and Prestidge,R.
TITLE Compounds and methods for treatment and diagnosis of mycobacterial
infections
JOURNAL Patent: US 5985287-A 83 16-NOV-1999;
FEATURES
source 1.20
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 4
LOCUS AR087520 20 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from patent US 5986084.
ACCESSION AR087520
VERSION AR087520.1 GI:10014283
KEYWORDS
SOURCE Unknown.
ORGANISM Unassigned.
REFERENCE 1 (bases 1 to 20)
AUTHORS Birsch,S., Weiss,P.A. and Jenny,L.
TITLE Ribonucleoside-derivative and method for preparing the same
JOURNAL Patent: US 5986084-A 1 16-NOV-1999;
FEATURES
source 1.20
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 5
LOCUS AR093312 20 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 83 from patent US 6001361.
ACCESSION AR093312
VERSION AR093312.1 GI:10020062
KEYWORDS
SOURCE Unknown.
ORGANISM Unassigned.
REFERENCE 1 (bases 1 to 20)
AUTHORS Tan,P., Hiyaama,J., Visser,E., Skinner,M., Scott,L. and Prestidge,R.
TITLE Mycobacterium vaccae antigens
JOURNAL Patent: US 6001361-A 83 14-DEC-1999;
FEATURES
source 1.20
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 6
LOCUS AR118970 20 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 96 from patent US 6150092.
ACCESSION AR118970
VERSION AR118970.1 GI:14100880
KEYWORDS
SOURCE Unknown.
ORGANISM Unassigned.
REFERENCE 1 (bases 1 to 20)
AUTHORS Uchida,K., Uchida,T., Tanaka,Y., Matsuda,Y. and Kondo,S.
TITLE Antisense nucleic acid compound targeted to VEGF
JOURNAL Patent: US 6150092-A 96 21-NOV-2000;
```

```
RESULT 4
LOCUS AR087520 20 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from patent US 5986084.
ACCESSION AR087520
VERSION AR087520.1 GI:10014283
KEYWORDS
SOURCE Unknown.
ORGANISM Unassigned.
REFERENCE 1 (bases 1 to 20)
AUTHORS Birsch,S., Weiss,P.A. and Jenny,L.
TITLE Ribonucleoside-derivative and method for preparing the same
JOURNAL Patent: US 5986084-A 1 16-NOV-1999;
FEATURES
source 1.20
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 5
LOCUS AR093312 20 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 83 from patent US 6001361.
ACCESSION AR093312
VERSION AR093312.1 GI:10020062
KEYWORDS
SOURCE Unknown.
ORGANISM Unassigned.
REFERENCE 1 (bases 1 to 20)
AUTHORS Tan,P., Hiyaama,J., Visser,E., Skinner,M., Scott,L. and Prestidge,R.
TITLE Mycobacterium vaccae antigens
JOURNAL Patent: US 6001361-A 83 14-DEC-1999;
FEATURES
source 1.20
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 6
LOCUS AR118970 20 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 96 from patent US 6150092.
ACCESSION AR118970
VERSION AR118970.1 GI:14100880
KEYWORDS
SOURCE Unknown.
ORGANISM Unassigned.
REFERENCE 1 (bases 1 to 20)
AUTHORS Uchida,K., Uchida,T., Tanaka,Y., Matsuda,Y. and Kondo,S.
TITLE Antisense nucleic acid compound targeted to VEGF
JOURNAL Patent: US 6150092-A 96 21-NOV-2000;
```

```

FEATURES                               Location/Qualifiers
source                                  1..20
                                         /organism="unknown"
                                         /mol_type="unassigned DNA"
ORIGIN
Query Match                           0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
    |||
    20 AAAAAAAAAAAAAAAAAA 1
    |||

RESULT 7
LOCUS AR121692 20 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 83 from patent US 6160093.
ACCESSION AR121692
VERSION AR121692.1 GI:14105268
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Visser, E.
TITLE Compounds and methods for treatment and diagnosis of mycobacterial
JOURNAL Patent: US 6160093-A 83 12-DEC-2000;
FEATURES Location/Qualifiers
source 1..20
      /organism="unknown"
      /mol_type="unassigned DNA"
ORIGIN
Query Match                           0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
    |||
    1 AAAAAAAAAAAAAAAAAA 20
    |||

Db 1 AAAAAAAAAAAAAAAAAA 20
    |||

RESULT 8
LOCUS AR123335 20 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6169176.
ACCESSION AR123335
VERSION AR123335.1 GI:14108301
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bruce, T.C. and Dev, A.P.
TITLE Deoxynucleic alkyl thioarea compounds and uses thereof
JOURNAL Patent: US 6169176-A 1 02-JAN-2001;
FEATURES Location/Qualifiers
source 1..20
      /organism="unknown"
      /mol_type="unassigned DNA"
ORIGIN
Query Match                           0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
    |||
    1 AAAAAAAAAAAAAAAAAA 20
    |||

Db 1 AAAAAAAAAAAAAAAAAA 20
    |||

```

```

RESULT 9
LOCUS AR141070/c 20 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6207819.
ACCESSION AR141070
VERSION AR141070.1 GI:14483566
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Manoharan, N. and Waier, M.A.
TITLE Compounds, processes and intermediates for synthesis of mixed
JOURNAL backbone oligomeric compounds
FEATURES Patent: US 6207819-A 1 27-MAR-2001;
source Location/Qualifiers
source 1..20
      /organism="unknown"
      /mol_type="unassigned DNA"
ORIGIN
Query Match                           0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
    |||
    20 AAAAAAAAAAAAAAAAAA 1
    |||

Db 20 AAAAAAAAAAAAAAAAAA 1
    |||

RESULT 10
LOCUS AR154115/c 20 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 14 from patent US 6238865.
ACCESSION AR154115
VERSION AR154115.1 GI:15122168
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Huang, Z. and Szostak, J.W.
TITLE Simple and efficient method to label and modify 3'-termini of RNA
JOURNAL using DNA polymerase and a synthetic template with defined overhang
FEATURES Patent: US 6238865-A 14 29-MAY-2001;
source Location/Qualifiers
source 1..20
      /organism="unknown"
      /mol_type="unassigned DNA"
ORIGIN
Query Match                           0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
    |||
    20 AAAAAAAAAAAAAAAAAA 1
    |||

Db 20 AAAAAAAAAAAAAAAAAA 1
    |||

RESULT 11
LOCUS AR164658 20 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 13 from patent US 6274321.
ACCESSION AR164658
VERSION AR164658.1 GI:16237754
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 20)

```

AUTHORS Blumberg,B.
TITLE High throughput functional screening of cDNAs
JOURNAL Patent: US 6274321-A 13 14-AUG-2001;
FEATURES Location/Qualifiers
SOURCE 1..20
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20
RESULT 12
E12676/c 20 bp DNA linear PAT 27-APR-1998
LOCUS E12676
DEFINITION Anti-HTLV-1 antisense oligonucleotide.
ACCESSION E12676
VERSION E12676.1 GI:3251508
KEYWORDS JP 1997052898-A/10.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 20)
AUTHORS Mizuguchi, M., Kurosaki, N., Makino, K., Koyanagi, Y. and Yamamoto, N.
TITLE ANTI-HTLV-1 ANTI-SENSE OLIGONUCLEOTIDE
JOURNAL Patent: JP 1997052898-A 10 25-FEB-1997;
SOYAKU GIJUTSU KENKYUSHO:KK
COMMENT OS None
OC Artificial sequences.
PN JP 1997052898-A/10
PD 25-FEB-1997
PF 09-AUG-1995 JP 1995224606
PI MIZUGUCHI MASATOSHI, KUROSAKI NAOKO, MAKINO KEISUKE, PI
KOYANAGI YOSHIO,
PI YAMAMOTO NAOKI
PC C07H21/04//A61K31/70;
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: Yes;
FH Key Location/Qualifiers
FH source 1..20
FT Location/Qualifiers
FEATURES source 1..20
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1799 AAAAAAAAAAAAAAAAAA 1818
Db 20 AAAAAAAAAAAAAAAAAA 1
RESULT 13
I36180 20 bp DNA linear PAT 13-MAY-1997
LOCUS I36180/c
DEFINITION Sequence 16 from patent US 5605662.
ACCESSION I36180
VERSION I36180.1 GI:2086693
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 20)
AUTHORS Heller, M.J. and Tu, E.
TITLE Active programmable electronic devices for molecular biological
JOURNAL Patent: US 5605662-A 16 25-FEB-1997;
FEATURES Location/Qualifiers
SOURCE 1..20
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1799 AAAAAAAAAAAAAAAAAA 1818
Db 20 AAAAAAAAAAAAAAAAAA 1
RESULT 14
AR213738 20 bp DNA linear PAT 25-SEP-2002
LOCUS AR213738
DEFINITION Sequence 83 from patent US 6406704.
ACCESSION AR213738
VERSION AR213738.1 GI:23311025
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 20)
AUTHORS Tan, P., Visser, E., Prestidge, R. and Watson, J.D.
TITLE Compounds and methods for treatment and diagnosis of mycobacterial
JOURNAL Patent: US 6406704-A 83 18-JUN-2002;
FEATURES Location/Qualifiers
SOURCE 1..20
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20
RESULT 15
AR222466 20 bp DNA linear PAT 26-SEP-2002
LOCUS AR222466
DEFINITION Sequence 26 from patent US 6429300.
ACCESSION AR222466
VERSION AR222466.1 GI:23329997
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 20)
AUTHORS Kurz, M., Lohse, P. and Wagner, R.
TITLE Peptide acceptor ligation methods
JOURNAL Patent: US 6429300-A 26 06-AUG-2002;
FEATURES Location/Qualifiers
SOURCE 1..20
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1799 AAAAAAAAAAAAAAAAAA 1818
|||
Db 1 AAAAAAAAAAAAAAAAAA 20
Search completed: March 7, 2004, 18:09:10
Job time : 10905 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 11:56:07 ; Search time 1110 Seconds
(Without alignments) 10785.065 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818
Sequence: 1 agagacacatcattatgta.....aaaaaaaaaaaaaaaaaa 2818

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1690386

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002s:*\n7: geneseqn2003as:*\n8: geneseqn2003bs:*\n9: geneseqn2003cs:*\n10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	0.7	20	AAQ25565	AAQ25565 Dye-coupled
2	20	0.7	20	AAQ33554	AAQ33554 Microstate
3	20	0.7	20	AAQ58578	AAQ58578 Sequence
4	20	0.7	20	AAQ94205	AAQ94205 Alpha-ano
5	20	0.7	20	AAQ75563	AAQ75563 Reverse t
6	20	0.7	20	AAQ75596	AAQ75596 Reverse t
7	20	0.7	20	AAQ90405	AAQ90405 T2 (synth
8	20	0.7	20	AAQ73704	AAQ73704 PCR prime
9	20	0.7	20	AAQ63649	AAQ63649 Anti-NTLV
10	20	0.7	20	AAV85582	AAV85582 LRPS PCR
11	20	0.7	20	AAV34591	AAV34591 M. vaccae
12	20	0.7	20	AAQ86606	AAQ86606 Oligonuc
13	20	0.7	20	AAQ37736	AAQ37736 Human mdm
14	20	0.7	20	AAQ37737	AAQ37737 Human mdm
15	20	0.7	20	AAQ27533	AAQ27533 Synthetic
16	20	0.7	20	AAQ27533	AAQ27533 Mycobacte
17	20	0.7	20	AAA40449	AAA40449 Electoche
18	20	0.7	20	AAA40448	AAA40448 Electoche
19	20	0.7	20	AAQ96372	AAQ96372 Primer us
20	20	0.7	20	AAQ35378	AAQ35378 Intersper
21	20	0.7	20	AAQ91117	AAQ91117 Oligonuc
22	20	0.7	20	AAA50193	AAA50193 2'-methox
23	20	0.7	20	AAQ14808	AAQ14808 Human gly

24	20	0.7	20	AAQ95176	AAQ95176 Human CDN
25	20	0.7	20	AAQ97238	AAQ97238 Phosphoro
26	20	0.7	20	AAQ87230	AAQ87230 Digoxigen
27	20	0.7	20	AAQ87241	AAQ87241 Poly T ol
28	20	0.7	20	AAQ10402	AAQ10402 DNA templ
29	20	0.7	20	AAQ16997	AAQ16997 Capture p
30	20	0.7	20	AAQ60896	AAQ60896 Conjugate
31	20	0.7	20	AAQ56348	AAQ56348 Oligonuc
32	20	0.7	20	AAQ28461	AAQ28461 Random ol
33	20	0.7	20	AAQ80891	AAQ80891 Human mdm
34	20	0.7	20	AAQ80890	AAQ80890 Human mdm
35	20	0.7	20	AAQ10371	AAQ10371 Oligonuc
36	20	0.7	20	AAQ99427	AAQ99427 Immunost
37	20	0.7	20	AAQ99099	AAQ99099 Immunost
38	20	0.7	20	AAQ99431	AAQ99431 Immunost
39	20	0.7	20	AAQ18246	AAQ18246 SNP speci
40	20	0.7	20	AAQ46465	AAQ46465 Oligonuc
41	20	0.7	20	AAQ78547	AAQ78547 Nucleotid
42	20	0.7	20	AAQ28351	AAQ28351 DNA oligo
43	20	0.7	20	AAQ01235	AAQ01235 Reverse p
44	20	0.7	20	AAQ29506	AAQ29506 Human mdm
45	20	0.7	20	AAQ29505	AAQ29505 Human mdm

ALIGNMENTS

AAQ25565/c	RESULT 1
AAQ25565 standard; DNA; 20 BP.	ID
AAQ25565;	AC
25-MAR-2003 (revised)	DT
02-DEC-1992 (first entry)	DT
Dye-coupled 3'-amino modified oligonucleotide.	XX
DNA synthesis; RNA; antisense strands; detection; ss.	XX
Synthetic.	OS
Key	XX
modified_base	XX
Location/Qualifiers	XX
20	XX
/*tag= a	XX
/note= "3-amino modified"	XX
EP490281-A1.	XX
17-JUN-1992.	XX
06-DEC-1991; 91BP-00120935.	XX
11-DEC-1990; 90DE-04039488.	XX
(FARH) HOECHST AG.	XX
Engels J, Herrlein M, Konrad R, Mag M;	XX
WPI; 1992-201578/25.	XX
New dye-coupled modified nucleosides, nucleotides and oligo:nucleotides -	XX
useful for synthesis of antisense DNA and RNA strands in presence of	XX
template, also for in-vivo and in-vitro detection of genetic material.	XX
Example; Page 9; 17pp; German.	XX
The sequence is an example of a dye coupled 3'-amino modified oligo-	XX
nucleotide, it can be used in the synthesis of DNA and RNA nucleosides,	XX
nucleotides and oligonucleotides and for the synthesis of opposite	XX
strands in the presence of a template strand and in fluorescence	XX
microscopic and microscopic detection in vivo and in vitro of genetic	XX
material. It is labelled with a fluorescent dye. See also AAQ25566 and	XX


```

FT misc_feature 1. 21
FT /tag= b
FT /note= "the glycosidic bonds between nucleotides are all
FT in the alpha-anomer form"
FT modified_base 20
FT /tag= a
FT /mod_base= OTHER
FT /note= "carries a group derived ffrom aminopropanediol"
XX
XX WO9429723-A1.
XX
XX 22-DEC-1994.
XX
XX 10-JUN-1994; 94WO-FR000689.
XX
XX 11-JUN-1993; 93FR-00007093.
XX
XX (CROS/) CROS P.
XX (KURF/) KURFURST R.
XX (BATT/) BATTAIL N.
XX (PIGA/) PIGA N.
XX
XX Cros P, Kurfurst R, Battail N, Piga N;
XX
XX WPI; 1995-036665/05.
XX
XX Assay device for hapten or its specific antibodies - comprises support
XX having competitive reagent immobilised via nucleic acid ligand to improve
XX orientation and accessibility.
XX
XX Example 1; Page 10; 39pp; French.
XX
XX Oligonucleotides (AAQ94201-094205) were synthesised for use as ligands.
XX The ligands are covalently linked to a hapten (esp. a steroid hormone) to
XX form a conjugate which is then immobilised on a solid support for
XX interaction with antibodies against the hapten. Nucleic acid ligands are
XX less likely to be recognised by the antibodies than are peptide ligands
XX and nucleic acids are also less likely to undergo intramolecular
XX organisation which interferes with accessibility of the hapten to the
XX antibodies. For immunodiagnosis of oestradiol, the active hapten
XX oestradiol-6-carboxymethoxime-N'-hydroxysuccinimide ester was used.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 20 BP; 0 A; 0 C; 0 G; 20 T; 0 U; 0 Other;
XX
XX Query Match 0.7%; Score 20; DB 2; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 1.5e+03;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1799 AAAAAAAAAAAAAAAAAA 1818
XX DB 20 AAAAAAAAAAAAAAAAAA 1
XX
XX RESULT 5
XX AAQ75563/c
XX ID AAQ75563 standard; DNA; 20 BP.
XX
XX AC AAQ75563;
XX
XX 04-AUG-1995 (first entry)
XX
XX Reverse transcription primer used in cDNA analysis technique.
XX
XX Analysis; gene expression; reverse transcription; primer; cDNA;
XX aggregate; restriction enzyme; ss.
XX
XX Synthetic.
XX
XX JP06303997-A.
XX
XX 01-NOV-1994.
XX

```

```

PF 16-APR-1993; 93JP-00112515.
XX
XX 16-APR-1993; 93JP-00112515.
XX
XX (NITE ) NIPPON TELEGRAPH & TELEPHONE CORP.
XX
XX WPI; 1995-018287/03.
XX
XX Analysis of cDNA and gene expression - by amplification of mRNA followed
XX by digestion with restriction enzymes.
XX
XX Disclosure; Page 5; 11pp; Japanese.
XX
XX A method for the analysis of cDNA comprises (a) preparing an aggregate of
XX double-stranded cDNAs by using an aggregate of mRNAs and a plural type of
XX labelled reverse transcription primers (GENESEQ files AAQ75547-075798)
XX and using the aggregate of mRNAs as the template for each reverse
XX transcription primer; (b) digesting each of the prepared aggregates of
XX the double-stranded cDNAs with restriction enzyme and; (c)
XX electrophoresing the digested aggregate of cDNAs in separate lanes. The
XX method can be used to analyse gene expression rapidly and easily
XX
XX Sequence 20 BP; 1 A; 0 C; 2 G; 17 T; 0 U; 0 Other;
XX
XX Query Match 0.7%; Score 20; DB 2; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 1.5e+03;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1796 CTCAAAAAAAAAAAAAAAAA 1815
XX DB 20 CTCAAAAAAAAAAAAAAAAA 1
XX
XX RESULT 6
XX AAQ75596/c
XX ID AAQ75596 standard; DNA; 20 BP.
XX
XX AC AAQ75596;
XX
XX 04-AUG-1995 (first entry)
XX
XX Reverse transcription primer used in cDNA analysis technique.
XX
XX Analysis; gene expression; reverse transcription; primer; cDNA;
XX aggregate; restriction enzyme; ss.
XX
XX Synthetic.
XX
XX JP06303997-A.
XX
XX 01-NOV-1994.
XX
XX 16-APR-1993; 93JP-00112515.
XX
XX 16-APR-1993; 93JP-00112515.
XX
XX (NITE ) NIPPON TELEGRAPH & TELEPHONE CORP.
XX
XX WPI; 1995-018287/03.
XX
XX Analysis of cDNA and gene expression - by amplification of mRNA followed
XX by digestion with restriction enzymes.
XX
XX Disclosure; Page 5; 11pp; Japanese.
XX
XX A method for the analysis of cDNA comprises (a) preparing an aggregate of
XX double-stranded cDNAs by using an aggregate of mRNAs and a plural type of
XX labelled reverse transcription primers (GENESEQ files AAQ75547-075798)
XX and using the aggregate of mRNAs as the template for each reverse
XX transcription primer; (b) digesting each of the prepared aggregates of
XX the double-stranded cDNAs with restriction enzyme and; (c)
XX electrophoresing the digested aggregate of cDNAs in separate lanes. The
XX method can be used to analyse gene expression rapidly and easily
XX

```

XX Sequence 20 BP; 2 A; 1 C; 0 G; 17 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2792 TTTGAAAAAAAAAAAAAAAA 2811
DB 20 TTTGAAAAAAAAAAAAAAAA 1

RESULT 7
AAQ90405/c
ID AAQ90405 standard; DNA; 20 BP.

AC AAQ90405;
XX
DT 08-JAN-1996 (first entry)

XX T2 (synthetic DNA probe with 5' amino terminal #4).

XX T2; HLA; dQa; self-addressable electronic device; SAED; hybridisation;
XX ss.

OS Synthetic.

XX Key Location/Qualifiers
FH misc_feature 1

FT /*tag= a
FT /note= "3' aminolink2 Thymine; allows binding to any
FT amine"

XX WO9512808-A1.

XX PD 11-MAY-1995.

XX PF 26-OCT-1994; 94MO-US012270.

XX PR 01-NOV-1993; 93US-00146504.

XX PA (NANO-) NANOGEN INC.

XX PI Heller MJ, Tu E;

XX DR WPI; 1995-185870/24.

XX New self-addressable electronic devices - used for multi-step and
PT multiplex reactions such as DNA hybridisation(s), clinical diagnostics
PT and bio/polymer synthesis.

XX Example 1; Page 41; 86pp; English.

XX The sequences represented by, AAQ90402-15 are synthetic DNA probes
CC containing 5' amino termini. The sequences shown in AAQ90390-401 are
CC synthetic DNA probes with 3' ribonucleoside termini. These sequences were
CC specific for the polymorphisms of HLA gene dQa. The sequences were used
CC in the device of the invention. This is a self-addressable electronic
CC device (SAED) that can be used to carry out multi-step and multiplex
CC reactions, such as nucleic acid hybridisations. The advantages of this
CC method are that these reactions can be carried out with complete and
CC precise electronic control, and that the rate, specificity and
CC sensitivity of these reactions are greatly improved at micro-locations

XX Sequence 20 BP; 0 A; 0 C; 0 G; 20 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 8

AA73704/c
ID AA73704 standard; DNA; 20 BP.

XX AA73704;

XX 27-FEB-1998 (first entry)

DE PCR primer used to prepare probes for diagnosing Alzheimer's disease.

XX PCR primer BK33; Alzheimer's disease; probe; diagnosis; fluorochrome;
KM yeast artificial chromosome library; YAC; chromosome 14; presentie; ss.

XX Synthetic.

XX FR2742758-A1.

XX 27-JUN-1997.

XX 28-OCT-1994; 94FR-00012941.

XX 28-OCT-1994; 94FR-00012941.

XX (ASFR-) ASSOC FR CONTRA MYOPATHIES ASSOC LOT.

XX Weissenbach J, Hellig R;

XX WPI; 1997-353201/33.

PT Probes for diagnosing Alzheimer's disease - hybridising with chromosome
PT 14 segments cloned in yeast artificial chromosome library.

PS Example 1; Page 8; 21pp; French.

XX PCR primers AAT73703-4 were used to prepare probes (containing Alu
CC repeats) for detecting a mutation in the locus of chromosome 14
CC associated with a presenile form of Alzheimer's disease. Each of the
CC probes hybridises with one of the two human chromosomal DNA segments
CC cloned in the CEPH yeast artificial chromosome (YAC) library under the
CC accession numbers YAC 93443 identifiable by genetic marker D14S251 and
CC YAC 854F5 (identifiable by genetic marker D14S76). The probes are useful
CC for diagnosis of the form of Alzheimer's disease associated with
CC chromosome 14 by a method comprising making a preparation of metaphase
CC chromosomes from the patient's lymphoblastoid cells on a microscope
CC slide, contacting the preparation under DNA hybridisation conditions with
CC the pair of probes or with one of the probes and another probe that
CC hybridises with YAC 905C2 from the same library, and detecting the
CC hybridised probes and their relative positions on a significant number of
CC pairs of chromosomes

XX Sequence 20 BP; 4 A; 4 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1844 GCTGACGCGCTGTATCCGAG 1863
DB 20 GCTGACGCGCTGTATCCGAG 1

RESULT 9

AAT63648/c
ID AAT63648 standard; DNA; 20 BP.

XX AAT63649;

XX 06-JUN-1997 (first entry)

DE Anti-HTLV antisense reference oligonucleotide HT.

KW antisense; complementary; tax gene; inhibit; HTLV-1;
 KW human T-cell lymphotropic virus type 1; viral antigen expression; ss.
 XX Synthetic.
 OS
 XX JP09052898-A.
 PN
 XX 25-FEB-1997.
 PD
 XX 09-AUG-1995; 95UP-00224606.
 PF
 XX 09-AUG-1995; 95UP-00224606.
 PR
 XX (SOYA-) SOYAKU GIUTSU KENKUSHO KK.
 PA
 XX WPI; 1997-197252/18.
 DR
 XX Anti-HTLV-1 anti-sense oligo:nucleotide - is complementary to region of
 PT tax gene from human T-cell lymphotropic virus type 1 and inhibits viral
 PT antigen expression.
 PT
 XX Example 1; Page 8; 10pp; Japanese.
 PS
 XX Oligonucleotides having a partial sequence consisting of at least 15
 CC bases of AAT63641 (an antisense oligo complementary to a region of the
 CC tax gene which can inhibit human T-cell lymphotropic virus type 1 (HTLV-
 CC 1) viral antigen expression) are claimed. In an example, six antisense
 CC oligos were designed, T1-T6 (AAT63650-55) and were compared to six oligos
 CC derived from other regions of HTLV-1, i.e. S1 (splice junction), P1
 CC (p21), R1 (rex), R1 (rex response element), B1 (env) and G1 (gag), four
 CC reference oligonucleotides T1S (tax-sense), HC (dc20), HT (dt20)
 CC (AAT63647-49) and a random 20mer (RAM) in a HTLV-1 virus antigen
 CC expression inhibiting test. Oligonucleotide T1 gave the best results
 SQ Sequence 20 BP; 0 A; 0 C; 0 G; 20 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1799 AAAAAAAAAAAAAAAAAA 1818
 Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 10
 AAV85582/C
 ID AAV85582 standard; DNA; 20 BP.
 XX
 AC AAV85582;
 XX
 AC 10-FEB-1999 (first entry)
 DT
 XX
 DE LRP5 PCR primer Gp1 1P.
 XX

KW LRP5; LDL-receptor related protein; LRP-3; IDDM; diagnosis; endocytosis;
 KW insulin dependent diabetes mellitus; autoimmune disease;
 KW glomerulonephritis; inflammation; viral infection; osteoporosis;
 KW hypercholesterolemia; Alzheimer's disease; low density lipoprotein;
 KW PCR primer; ss.
 XX
 XX Synthetic.
 OS
 XX Homo sapiens.
 XX
 PV WO9846743-A1.
 XX
 PD 22-OCT-1998.
 XX
 PF 15-APR-1998; 98WO-GB001102.
 XX
 PR 15-APR-1997; 97US-0043553P.
 XX
 PR 05-JUN-1997; 97US-0048740P.
 XX

PA (WELL) WELLCOME TRUST LTD.
 XX (MERT) MERCK & CO INC.
 XX

PI Todd JA, Hess JW, Caskey CT, Cox RD, Gerhold D, Hammond H;
 PI Hey P, Kawaguchi Y, Merriman TR, Metzker ML, Nakagawa Y;
 PI Phillips MS, Twells RCJ;
 XX

DR WPI; 1998-594573/50.

PT New isolated LDL-receptor related protein - used to develop products for
 PT treating, e.g. elevated triglyceride levels, diabetes, autoimmune
 PT disorder, inflammation or Alzheimer's disease.
 PS

PS Claim 12; Page 98; 200pp; English.

CC The present invention describes LRP5 (low density lipoprotein (LDL)
 CC receptor related protein, previously designated LRP-3). AAV85552 to
 CC AAV85586 represent PCR primer used for obtaining LRP5 cDNA. Nucleic acid
 CC molecules (NAs) encoding LRP5 can be used for determining if an
 CC individual is susceptible to insulin dependent diabetes mellitus (IDDM).
 CC The NAs or proteins can be used for reducing triglyceride levels in the
 CC serum of an individual. Therapies that affect LRP5 may also be useful in
 CC the treatment of autoimmune diseases such as glomerulonephritis, diseases
 CC and disorders involving disruption of endocytosis and/or antigen
 CC presentation, cytokine clearance and/or inflammation, viral infection,
 CC pathogenic bacterial toxin contamination, elevation of free fatty acids
 CC or hypercholesterolemia, type 2 diabetes, osteoporosis, Alzheimer's
 CC disease and cardiovascular disease. Products from the present invention
 CC can also be used for detection, diagnosis and drug screening
 SQ Sequence 20 BP; 4 A; 6 C; 5 G; 5 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1901 CAGGAGTTGAGACCGCCT 1920
 Db 20 CAGGAGTTGAGACCGCCT 1

RESULT 11
 AAV34591
 ID AAV34591 standard; DNA; 20 BP.
 XX
 AC AAV34591;
 XX

DE 25-AUG-1998 (first entry)
 DT
 XX
 DE M. vaccae antigenic sequence hybridising oligo AD12.
 XX

KW Mycobacterium vaccae; antigen; therapy; prevention; cytokine production;
 KW M. avium; M. tuberculosis; immune response enhancer; cell proliferation;
 KW mycobacteria infection; vaccine; cancer; ss.
 XX

OS Synthetic.
 OS Mycobacterium vaccae.
 XX
 PN WO9808542-A2.
 XX

PD 05-MAR-1998.

PF 28-AUG-1997; 97WO-NZ000105.

PR 29-AUG-1996; 96US-00705347.

PR 12-JUN-1997; 97US-00873970.
 XX (GENE-) GENESIS RES & DEV CORP.

PI Tan P, Hiyama J, Visser E, Skinner MA, Scott LM, Prestidge RL;
 DR WPI; 1998-216926/19.

PT Mycobacterium vaccae polypeptides - used to develop products for use in
PT detection, therapy and prevention of mycobacteria infections or as immune
PT response enhancers.
PS
XX Example 8; Page 99; 153pp; English.
CC This oligonucleotide is used in the DNA cloning strategies of the
CC Mycobacterium vaccae antigens. The invention provides M. vaccae
CC polypeptides that comprise an immunogenic portion of a soluble M. vaccae
CC antigen, or a variant, where the antigen induces an immune response in
CC patients previously exposed to a mycobacterium. Such M. vaccae
CC polypeptides can be used in methods for enhancing non-specific immune
CC response. The methods and products can be used for the detection,
CC treatment and prevention of infectious diseases caused by mycobacteria
CC such as M. vaccae, M. avium or M. tuberculosis. The products also have
CC the ability to induce cell proliferation and cytokine production (e.g.
CC interferon-gamma and interleukin-12 production) in T cells, NK cells, B
CC cells, or macrophages. They can be used for enhancing immune responses
CC for use in vaccines or immunotherapy of infectious diseases and cancers
XX
SQ Sequence 20 BP; 20 A; 0 C; 0 G; 0 T; 0 U; 0 Other;
Query Match 0.7%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1799 AAAAAAAAAAAAAAAAAA 1818
DB 1 AAAAAAAAAAAAAAAAAA 20
RESULT 12
AAAT86606/c
ID AAAT86606 standard; DNA; 20 BP.
XX
AC AAAT86606;
XX
DT 04-JUN-1998 (first entry)
XX
DE Oligonucleotide separated by capillary affinity gel electrophoresis.
XX
KM Capillary affinity gel electrophoresis; separation; polymer-gel;
KM polyacrylamide; ss.
XX
OS Synthetic.
OS
PN WO9745721-A1.
XX
PD 04-DEC-1997.
XX
PF 23-MAY-1997; 97WO-EP002647.
XX
PR 24-MAY-1996; 96CH-00001320.
XX
PA (NOVS) NOVARTIS AG.
XX
PI Muscate A, Paulus A, Natt F;
XX
DR WPI; 1998-041763/04.
XX
PT Separation of electrically charged target molecules - by capillary
PT affinity gel electrophoresis using polymer-gel to which receptors for
PT target molecules are bound.
XX
PS Example D3; Page 25; 41pp; English.
XX
CC A mixture of oligonucleotides (AAAT86604-7) were separated by a new
CC process using capillary affinity gel electrophoresis. The invention
CC relates to selective separation of electrically charged target molecules
CC in an analytical mixture. It comprises capillary affinity gel
CC electrophoresis using a capillary tube which is at least partly filled
CC with a polymer gel. Receptors for target molecules are covalently bound
CC to the polymer. An electric field of at least 50 volts/cm is applied. The

CC capillary tube is charged with the analytical mixture. In a first
CC separation stage, the target molecules in the mixture are bound to the
CC receptors and the remaining components are eluted, optionally whilst
CC splitting open. In a second stage, the elution conditions are changed,
CC optionally in stages, so that the affinity of the target molecules for
CC the receptor is eliminated and the target molecules are eluted and
CC detected, optionally whilst splitting open. The process is useful for
CC selective separation and/or determination of charged organic compounds,
CC such as oligonucleotides, peptides or carbohydrates. It may be used, e.g.
CC for isolation of specific proteins and DNA molecules, purification of
CC antibodies, analysis of antisense compounds or screening for enzyme
CC inhibitors. The process achieves higher resolution and selectivity than
CC prior art processes, especially in the case of complex biological
CC analytical mixtures. It has high sensitivity, even with small amounts of
CC samples. The derivatised polymers may be synthesised specifically using
CC standard methods
XX
SQ Sequence 20 BP; 0 A; 0 C; 0 G; 20 T; 0 U; 0 Other;
Query Match 0.7%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1799 AAAAAAAAAAAAAAAAAA 1818
DB 20 AAAAAAAAAAAAAAAAAA 1
RESULT 13
AAZ37736
ID AAZ37736 standard; DNA; 20 BP.
XX
AC AAZ37736;
XX
DT 07-JAN-2000 (first entry)
XX
DE Human mdm2 phosphorothioate oligodeoxynucleotide #266.
XX
KM Human mdm2 gene; proliferation; tumour; phosphorothioate; p53; cancer;
KM antisense; modulation; oligonucleotide; expression; inhibition;
KM hyperproliferation; blood cancer; brain cancer; breast cancer;
KM lung cancer; soft tissue cancer; psoriasis; fibrosis; atherosclerosis;
KM restenosis; ss.
XX
OS Synthetic.
OS
OS Homo sapiens.
XX
XX WO9949065-A1.
XX
XX 30-SEP-1999.
XX
XX 26-MAR-1999; 99WO-US006702.
XX
XX 26-MAR-1998; 98US-00048810.
XX
XX (ISIS-) ISIS PHARM INC.
XX
PI Miraglia LJ, Nero P, Graham MJ, Cowse LM;
XX
XX WPI; 1999-610754/52.
XX
XX New antisense compounds used to treat eg. hyperproliferative conditions.
XX
PS Example 9; Page 55; 157pp; English.
XX
CC AAZ37473-Z37738 represent human mdm2 phosphorothioate oligonucleotides.
CC AAZ37471, AAZ37472, AAZ37739, AAZ37740 and AAZ37741 are used in the
CC exemplification of the present invention. The present invention describes
CC novel nucleotide antisense compounds, targeted to the 5' untranslated
CC translation termination codon, or 3' untranslated region of a nucleic
CC acid encoding human mdm2, that modulates expression of human mdm2. The
CC oligonucleotides mediate their effect by antisense inhibition of
CC hyperproliferative gene expression. The antisense compound is used to

CC treat an animal having a disease or condition associated with mdm2.
 CC particularly a hyperproliferative condition, more particularly cancer,
 CC especially of the blood, brain, breast, lung or soft tissue, or
 CC psoriasis, fibrosis, atherosclerosis or restenosis
 XX
 SQ Sequence 20 BP; 3 A; 7 C; 6 G; 4 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TCCGACGACTTTGGAGGCC 1877
 |||||
 1 TCCGACGACTTTGGAGGCC 20

RESULT 14
 AA237737
 ID AA237737 standard; DNA; 20 BP.

AC AA237737;
 DT 07-JAN-2000 (first entry)

DE Human mdm2 phosphorothioate oligodeoxynucleotide #267.

KW Human mdm2 gene; proliferation; tumour; phosphorothioate; p53; cancer;
 KW antisense; modulation; oligonucleotide; expression; inhibition;
 KW hyperproliferation; blood cancer; brain cancer; breast cancer;
 KW lung cancer; soft tissue cancer; psoriasis; fibrosis; atherosclerosis;
 KW restenosis; ss.

XX Synthetic.
 OS Homo sapiens.

PN WO9949065-A1.

PD 30-SEP-1999.

PF 26-MAR-1999; 99WO-US006702.

PR 26-MAR-1999; 98US-00048810.

PA (ISIS-) ISIS PHARM INC.

PI Miraglia LJ, Nero P, Graham MJ, Monia BP, Cowest LM;

DR WPI; 1999-610754/52.

PT New antisense compounds used to treat eg. hyperproliferative conditions.
 PS Example 9; Page 55; 157pp; English.

CC AA237473-227738 represent human mdm2 phosphorothioate oligonucleotides.
 CC AA237471, AA237472, AA237739, AA237740 and AA237741 are used in the
 CC exemplification of the present invention. The present invention describes
 CC novel nucleotide antisense compounds, targeted to the 5' untranslated,
 CC translation termination codon, or 3' untranslated region of a nucleic
 CC acid encoding human mdm2, that modulates expression of human mdm2. The
 CC oligonucleotides mediate their effect by antisense inhibition of
 CC hyperproliferative gene expression. The antisense compound is used to
 CC treat an animal having a disease or condition associated with mdm2,
 CC particularly a hyperproliferative condition, more particularly cancer,
 CC especially of the blood, brain, breast, lung or soft tissue, or
 CC psoriasis, fibrosis, atherosclerosis or restenosis
 CC
 SQ Sequence 20 BP; 4 A; 7 C; 3 G; 6 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1851 CCTGATATCCGACGACTTTG 1870

Db |||||
 1 CCTGATATCCGACGACTTTG 20

RESULT 15
 AAX27533/c
 ID AAX27533 standard; RNA; 20 BP.

AC AAX27533;

DT 27-MAY-1999 (first entry)

DE Synthetic RNA sequence produced by the method of the invention.

KW silyloxymethyl; phosphonate; silyloxymethyl halide; diagnosis; ss;
 KW cyanoethyl phosphoramidate coupling; isomerisation; steric hindrance.

OS Synthetic.

PN WO9909044-A1.

PD 25-FEB-1999.

PF 17-AUG-1998; 98WO-EP005215.

PR 18-AUG-1997; 97CH-00001931.

PA (PITS/) PITSCH S.
 PA (WEIS/) WEISS P.A.
 PA (JENN/) JENNY L.

PI Pitsch S, Weiss PA, Jenny L;

DR WPI; 1999-180963/15.

PT 2-Silyloxymethyl ribonucleosides and their phosphonate derivatives - have
 PT high purity, use in machine synthesis of ribonucleic acids, enable longer
 PT oligonucleotide chain construction, and larger amounts.

PS Example 6; Page 25; 38pp; English.

CC The invention relates to silyloxymethyl protected D- or L-ribonucleosides
 CC and their phosphonates (I), and silyloxymethyl halides (II). (I) are
 CC intermediates for synthesis of RNA-oligonucleotides with predetermined
 CC nucleotide sequence, particularly by machine synthesis. The groups
 CC specified above, apart from those on silyl, are those particularly for
 CC the cyanoethyl phosphoramidate coupling. Uses of the oligonucleotide
 CC products in diagnosis, therapy, and as research tools, are well known,
 CC and are not dealt with in detail. (II) is an intermediate for (I). The
 CC silyloxymethyl halide reagent is easy to prepare, and yields are high.
 CC Introduction of the silyloxymethyl group into the ribonucleoside is
 CC simple and rapid, and the acetal bond formed does not migrate,
 CC eliminating particularly the prior art problem of 2' to 3' isomerisation.
 CC The methylatedoxy group spacer between the silyl group and nucleoside
 CC ring results in less steric hindrance than bulky direct silyloxy
 CC linkages, enabling first, a range of choices for the silyl substituents,
 CC to provide, e.g., acid or base stability; and second, higher yields in
 CC coupling. Purer products are therefore obtained than in prior art,
 CC enabling larger quantities and longer chains of oligoribonucleotides to
 CC be synthesised successfully, and in shorter times

SQ Sequence 20 BP; 0 A; 0 C; 0 G; 0 T; 20 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
 |||||
 Db 20 AAAAAAAAAAAAAAAAAA 1

Search completed: March 7, 2004, 15:07:19

Mon Mar 8 10:28:42 2004

us-09-966-880a-7.Oligo.rng

Page 8

Job time : 113 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic acid search, using sw model

Run on: March 7, 2004, 14:32:38 ; Search time 216 Seconds

(without alignments)
7240.054 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818
Sequence: 1 agagaccatcatatga.....aaaaaaaaaaaaaaaaaa 2818

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 353258

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Listing first 45 summaries

Database :

Issued_Patents.NA.*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/6C.COMB.seq:*
6: /cgn2_6/prodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	0.7	20	1 US-08-146-504-16	Sequence 16, Appl
C 2	20	0.7	20	2 US-08-379-593-5	Sequence 5, Appl
C 3	20	0.7	20	2 US-08-725-976-16	Sequence 16, Appl
C 4	20	0.7	20	2 US-08-997-080-83	Sequence 83, Appl
C 5	20	0.7	20	2 US-08-997-362-83	Sequence 83, Appl
C 6	20	0.7	20	2 US-08-965-780-1	Sequence 1, Appl
C 7	20	0.7	20	2 US-08-873-970-83	Sequence 83, Appl
C 8	20	0.7	20	3 US-08-765-340-96	Sequence 83, Appl
C 9	20	0.7	20	3 US-09-095-855-83	Sequence 83, Appl
C 10	20	0.7	20	3 US-09-407-675-1	Sequence 1, Appl
C 11	20	0.7	20	3 US-09-280-805-266	Sequence 266, App
C 12	20	0.7	20	3 US-09-280-805-267	Sequence 267, App
C 13	20	0.7	20	3 US-09-250-075-1	Sequence 1, Appl
C 14	20	0.7	20	3 US-09-173-936B-14	Sequence 14, Appl
C 15	20	0.7	20	3 US-09-454-704A-13	Sequence 13, Appl
C 16	20	0.7	20	4 US-09-488-856A-62	Sequence 62, Appl
C 17	20	0.7	20	4 US-09-324-542-83	Sequence 83, Appl
C 18	20	0.7	20	4 US-09-205-426-83	Sequence 83, Appl
C 19	20	0.7	20	4 US-09-619-103-26	Sequence 26, Appl
C 20	20	0.7	20	4 US-09-726-096A-1	Sequence 1, Appl
C 21	20	0.7	20	4 US-09-733-294A-82	Sequence 82, Appl
C 22	20	0.7	20	4 US-09-603-810-55	Sequence 55, Appl
C 23	20	0.7	20	4 US-09-976-978A-55	Sequence 55, Appl
C 24	20	0.7	20	4 US-09-060-299-78	Sequence 78, Appl
C 25	20	0.7	20	4 US-09-402-923A-78	Sequence 78, Appl
C 26	20	0.7	20	4 US-09-679-299A-70	Sequence 70, Appl
C 27	20	0.7	20	4 US-09-344-260A-10	Sequence 10, Appl

28	20	0.7	20	4 US-09-961-949A-55	Sequence 55, Appl
29	20	0.7	20	4 US-09-966-491A-55	Sequence 55, Appl
30	20	0.7	20	4 US-09-957-313A-55	Sequence 55, Appl
31	20	0.7	20	4 US-09-966-312-55	Sequence 55, Appl
32	20	0.7	20	4 US-09-975-062A-55	Sequence 55, Appl
33	20	0.7	20	4 US-09-976-971A-55	Sequence 55, Appl
34	20	0.7	20	5 PCT-US93-07603-6	Sequence 6, Appl
35	19	0.7	19	1 US-08-629-939-10	Sequence 10, Appl
C 36	19	0.7	19	1 US-08-759-873-10	Sequence 10, Appl
C 37	19	0.7	19	1 US-08-756-728A-1	Sequence 1, Appl
C 38	19	0.7	19	2 US-08-469-852A-2	Sequence 2, Appl
C 39	19	0.7	19	3 US-08-271-882B-16	Sequence 16, Appl
C 40	19	0.7	19	3 US-08-295-509B-2	Sequence 2, Appl
C 41	19	0.7	19	3 US-09-234-237-1	Sequence 1, Appl
C 42	19	0.7	19	3 US-09-016-520-20	Sequence 20, Appl
C 43	19	0.7	19	3 US-09-016-520-21	Sequence 21, Appl
C 44	19	0.7	19	3 US-09-016-520-22	Sequence 22, Appl
C 45	19	0.7	19	3 US-09-016-520-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-08-146-504-16/c
Sequence 16, Application US/08146504
Patent No. 5605662
GENERAL INFORMATION:
APPLICANT: Heller, Michael J., and Tu, Eugene
TITLE OF INVENTION: SELF-ADDRESSABLE SELF-ASSEMBLING
TITLE OF INVENTION: MICROELECTRONIC SYSTEMS AND DEVICES FOR
TITLE OF INVENTION: MOLECULAR BIOLOGICAL ANALYSIS AND
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 611 West Sixth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM compatible
OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
FILING DATE: No. 5605662member 1, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 203/218
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 673510
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 20
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-146-504-16
Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818

DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 2

US-08-379-593-5/c

Sequence 5, Application US/08379593

Patent No. 5849480

GENERAL INFORMATION:

APPLICANT: Cros, Philippe

APPLICANT: Kufirst, Robin

APPLICANT: Battail, Nicole

TITLE OF INVENTION: HAPTEN ASSAY DEVICE AND USE THEREOF

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESSES:

ADDRESSEE: OLIF & BERRIDGE

STREET: 700 South Washington Street, Suite 300

CITY: Alexandria

STATE: Virginia

COUNTRY: USA

ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Floppy disk, 1.44M storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/379,593

FILING DATE: 02-FEB-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36056

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-836-6400

TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "SYNTHETIC DNA"

FEATURE:

OTHER INFORMATION: consists of nucleosides with an alpha anomer and carries

US-08-379-593-5

QY 1799 AAAAAAAAAAAAAAAAAA 1818

DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 3

US-08-725-976-16/c

Sequence 16, Application US/08725976

Patent No. 5929208

GENERAL INFORMATION:

APPLICANT: Heller, Michael J. and Tu, Eugene

TITLE OF INVENTION: METHODS FOR ELECTRONIC SYNTHESIS OF POLYMERS

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 MB

COMPUTER: IBM compatible

OPERATING SYSTEM: WINDOWS (VERSION 3.0)

SOFTWARE: WordPerfect (Version 6.0)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/725,976

FILING DATE: October 4, 1996

CLASSIFICATION: 422

PRIOR APPLICATION DATA: including application

PRIOR APPLICATION DATA: described below:

APPLICATION NUMBER: 08/146,504

FILING DATE: No. 5929208ember 1, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Murphy, David B.

REGISTRATION NUMBER: 31,125

REFERENCE/DOCKET NUMBER: 222/211

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 20

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-725-976-16

QY 1799 AAAAAAAAAAAAAAAAAA 1818

DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 4

US-08-997-080-83

Sequence 83, Application US/08997080

Patent No. 5968524

GENERAL INFORMATION:

APPLICANT: WATSON, JAMES D.

APPLICANT: TAN, PAUL L.J.

TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-

NUMBER OF SEQUENCES: 194

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Law Offices of Ann W. Speckman

STREET: 2601 Elliott Avenue, Suite 4185

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/997,080

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Sleath, Janet

REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other
US-08-997-080-83

Query Match 0.7%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
DB 1 AAAAAAAAAAAAAAAAAA 20

RESULT 5
US-08-997-362-83
Sequence 83, Application US/08997362
Patent No. 5985287
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: HiYama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997.362
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873.970
FILING DATE: June 12, 1997
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705.347
FILING DATE: August 29, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleach, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: Other
US-08-997-362-83

Query Match 0.7%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
DB 1 AAAAAAAAAAAAAAAAAA 20

RESULT 6
US-08-965-780-1/c
Sequence 1, Application US/08965780
Patent No. 5986084
GENERAL INFORMATION:
APPLICANT: Pitsch, Stefan
APPLICANT: Weiss, Patrick A.
APPLICANT: Jenny, Luzi
TITLE OF INVENTION: RIBONUCLEOSIDE-DERIVATIVE AND METHOD FOR
PREPARING THE SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: KUBOVCIK & KUBOVCIK
STREET: 900 17th Street, N.W., Suite 990
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965.780
FILING DATE: 07-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 01931/97
FILING DATE: 18-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Kubovcik, Ronald J.
REGISTRATION NUMBER: 25,401
REFERENCE/DOCKET NUMBER: FREI-002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-9023
TELEFAX: 202-887-9093
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "oligoribonucleotide"
US-08-965-780-1

Query Match 0.7%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 7
US-08-873-970-83
Sequence 83, Application US/08873970
Patent No. 6001361
GENERAL INFORMATION:

APPLICANT: Tan, Paul
APPLICANT: Hiwama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
City: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/873,970
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other
US-08-873-970-83

Query Match 0.7%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
1 AAAAAAAAAAAAAAAAAA 20
Db

RESULT 8
US-08-765-340-96/C
Sequence 96, Application US/08765340
Patent No. 6150092
GENERAL INFORMATION:
APPLICANT: UCHIDA, K.,
APPLICANT: UCHIDA, T.,
APPLICANT: TANAKA, Y.,
APPLICANT: MATSUDA, Y.,
APPLICANT: KONDO, S.,
TITLE OF INVENTION: AN ANTISENSE NUCLEIC ACID
TITLE OF INVENTION: COMPOUND
NUMBER OF SEQUENCES: 185
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA

ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version
SOFTWARE: #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/765,340
FILING DATE: 23-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 145146/94
FILING DATE: 27-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 311130/94
FILING DATE: 21-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: SERUNIAN, LESLIE
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 1452-4005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "synthetic DNA"
US-08-765-340-96

Query Match 0.7%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
20 AAAAAAAAAAAAAAAAAA 1
Db

RESULT 9
US-09-095-855-83
Sequence 83, Application US/09095855
Patent No. 6160093
GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
Treatment and Diagnosis of Mycobacterial Infections
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996

APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other
US-09-955-85
Query Match 0.7%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1799 AAAAAAAAAAAAAAAAAA 1818
1 AAAAAAAAAAAAAAAAAA 20
Db 1 AAAAAAAAAAAAAAAAAA 20
RESULT 10
US-09-407-675-1
Sequence 1, Application US/09407675
Patent No. 6169176
GENERAL INFORMATION:
APPLICANT: Bruice, Thomas C.
APPLICANT: Arya, Dev P.
TITLE OF INVENTION: DEOXYNUCLEIC ALKYL THIUREA COMPOUNDS AND USES THEREOF
FILE REFERENCE: 30448.65US02
CURRENT APPLICATION NUMBER: US/09/407,675
CURRENT FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: 09/347,443
PRIOR FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: 60/091,481
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/111,800
PRIOR FILING DATE: 1998-12-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence: Oligo 1
US-09-407-675-1
Query Match 0.7%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1799 AAAAAAAAAAAAAAAAAA 1818
1 AAAAAAAAAAAAAAAAAA 20
Db 1 AAAAAAAAAAAAAAAAAA 20
RESULT 11
US-09-280-805-266
Sequence 266, Application US/09280805
Patent No. 6184212
GENERAL INFORMATION:
APPLICANT: Loren J. Miraglia, Pamela Nero, Mark J.
APPLICANT: Graham, Brett P. Monia

TITLE OF INVENTION: ANTISENSE MODULATION OF HUMAN MDM2
NUMBER OF SEQUENCES: 271
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Jane Massey Licata
STREET: 66 East Main Street
CITY: Marlton
STATE: NJ
COUNTRY: U.S.A.
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PC
OPERATING SYSTEM: WINDOWS 95
SOFTWARE: WORDPERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/280,805
FILING DATE: herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/048,810
FILING DATE: March 26, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Licata, Jane Massey
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0346
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-810-1515
TELEFAX: 609-810-1454
INFORMATION FOR SEQ ID NO: 266:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: single
TOPOLOGY: linear
ANTI-SENSE: Yes
US-09-280-805-266
Query Match 0.7%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1858 TCCGAGCACTTGGAGGCC 1877
1 TCCGAGCACTTGGAGGCC 20
Db 1 TCCGAGCACTTGGAGGCC 20
RESULT 12
US-09-280-805-267
Sequence 267, Application US/09280805
Patent No. 6184212
GENERAL INFORMATION:
APPLICANT: Loren J. Miraglia, Pamela Nero, Mark J.
APPLICANT: Graham, Brett P. Monia
TITLE OF INVENTION: ANTISENSE MODULATION OF HUMAN MDM2
NUMBER OF SEQUENCES: 271
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Jane Massey Licata
STREET: 66 East Main Street
CITY: Marlton
STATE: NJ
COUNTRY: U.S.A.
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PC
OPERATING SYSTEM: WINDOWS 95
SOFTWARE: WORDPERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/280,805
FILING DATE: herewith
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/048,810
FILING DATE: March 26, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Licata, Jane Massey
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0346
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-810-1515
TELEFAX: 609-810-1454
INFORMATION FOR SEQ ID NO: 267:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: Yes
US-09-280-805-267

Query Match 0.7%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1851 CCTGTATCCGACACTTGG 1870
DB 1 CCTGTATCCGACACTTGG 20

RESULT 13
US-09-250-075-1/c
Sequence 1, Application US/09250075
Patent No. 6207819
GENERAL INFORMATION:
APPLICANT: Manoharan, Muthiah
APPLICANT: Maier, Martin A
TITLE OF INVENTION: Compounds Processes And Intermediates For Synthesis Of
TITLE OF INVENTION: Mixed Backbone Oligomeric Compounds
FILE REFERENCE: ISIS3299
CURRENT APPLICATION NUMBER: US/09/250,075
CURRENT FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentn Ver. 2.1
SEQ ID NO 1
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(19)
OTHER INFORMATION: 2'-methoxyethoxy (MOE)
OTHER INFORMATION: Description of Artificial Sequence: No. 620781.9e1
OTHER INFORMATION: Sequence
US-09-250-075-1

Query Match 0.7%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 14
US-09-173-936B-14/c
Sequence 14, Application US/09173936B
Patent No. 6238865
GENERAL INFORMATION:
APPLICANT: Zhen, Huang; Szostak, Jack W.
TITLE OF INVENTION: A Simple and Efficient Method to Label and Modify 3'-
Terminal
of RNA Using DNA Polymerase and a Synthetic Template with D
Nucleotides

NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cohen, Pontani, Lieberman & Pavana
STREET: 551 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10176
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 Inch Diskette
COMPUTER: IBM-MS
OPERATING SYSTEM: Window 95
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/173,936B
FILING DATE: 16-Oct-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/063,757
FILING DATE: 17-Oct-1997
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-173-936B-14

Query Match 0.7%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 15
US-09-454-704A-13
Sequence 13, Application US/09454704A
Patent No. 6274321
GENERAL INFORMATION:
APPLICANT: Blumberg, Bruce
TITLE OF INVENTION: High Throughput Functional Screening of
TITLE OF INVENTION: CDNs
FILE REFERENCE: P-UC 3662
CURRENT APPLICATION NUMBER: US/09/454,704A
CURRENT FILING DATE: 1999-12-03
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: cDNA
US-09-454-704A-13

Query Match 0.7%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
DB 1 AAAAAAAAAAAAAAAAAA 20

Search completed: March 7, 2004, 20:11:12
Job time : 218 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 18:09:14 ; Search time 963 Seconds

(without alignments)
10702.641 Million cell updates/sec

Title: US-09-966-880A-7

Sequence: 1 agagagaccctatcatatga.....aaaaaaaaaaaaaaaa 2818

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2421054 seqs, 1828716029 residues

Word size : 0

Total number of hits satisfying chosen parameters: 452776

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-Processing: Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	0.7	20	US-09-752-983-266	Sequence 266, App
2	20	0.7	20	US-09-752-983-267	Sequence 267, App
3	20	0.7	20	US-09-733-294A-82	Sequence 82, App
4	20	0.7	20	US-09-973-788A-55	Sequence 55, App
5	20	0.7	20	US-09-973-638A-55	Sequence 55, App
6	20	0.7	20	US-09-974-007-55	Sequence 55, App
7	20	0.7	20	US-09-976-617A-55	Sequence 55, App
8	20	0.7	20	US-09-961-949A-55	Sequence 55, App
9	20	0.7	20	US-09-760-500A-55	Sequence 55, App
10	20	0.7	20	US-09-967-409A-55	Sequence 55, App
11	20	0.7	20	US-09-975-062A-55	Sequence 55, App
12	20	0.7	20	US-09-976-378A-55	Sequence 55, App
13	20	0.7	20	US-09-976-577-55	Sequence 55, App
14	20	0.7	20	US-09-771-554-5	Sequence 5, Appl
15	20	0.7	20	US-09-966-312-55	Sequence 55, Appl

16	20	0.7	20	US-09-927-777A-55	Sequence 55, Appl
17	20	0.7	20	US-09-927-777A-70	Sequence 70, Appl
18	20	0.7	20	US-09-966-491A-55	Sequence 55, Appl
19	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
20	20	0.7	20	US-09-976-971A-4	Sequence 4, Appl
21	20	0.7	20	US-09-853-753-4	Sequence 83, Appl
22	20	0.7	20	US-09-880-505-83	Sequence 55, Appl
23	20	0.7	20	US-09-820-272B-55	Sequence 55, Appl
24	20	0.7	20	US-09-888-326-2	Sequence 2, Appl
25	20	0.7	20	US-09-888-326-83	Sequence 838, App
26	20	0.7	20	US-09-888-326-83	Sequence 838, App
27	20	0.7	20	US-09-981-344-55	Sequence 55, Appl
28	20	0.7	20	US-09-957-318A-55	Sequence 55, Appl
29	20	0.7	20	US-09-974-500A-55	Sequence 55, Appl
30	20	0.7	20	US-09-975-376A-55	Sequence 55, Appl
31	20	0.7	20	US-09-957-313A-55	Sequence 55, Appl
32	20	0.7	20	US-09-912-014-16	Sequence 16, Appl
33	20	0.7	20	US-09-997-672-40	Sequence 40, Appl
34	20	0.7	20	US-09-976-862A-55	Sequence 55, Appl
35	20	0.7	20	US-09-881-535-2	Sequence 2, Appl
36	20	0.7	20	US-09-776-479-226	Sequence 226, App
37	20	0.7	20	US-09-776-479-556	Sequence 556, App
38	20	0.7	20	US-09-976-601A-55	Sequence 560, App
39	20	0.7	20	US-09-975-058A-55	Sequence 55, Appl
40	20	0.7	20	US-09-976-968A-55	Sequence 55, Appl
41	20	0.7	20	US-10-208-357-26	Sequence 55, Appl
42	20	0.7	20	US-10-051-643-83	Sequence 83, Appl
43	20	0.7	20	US-10-176-055-11	Sequence 11, Appl
44	20	0.7	20	US-10-117-267-1	Sequence 117, Appl
45	20	0.7	20	US-10-112-653-218	Sequence 218, App

ALIGNMENTS

RESULT 1
US-09-752-983-266
Sequence 266, Application US/09752983
Patent No. US20010016575A1
GENERAL INFORMATION:
APPLICANT: Loren J. Miraglia, Pamela Nero, Mark J.
APPLICANT: Graham, Brett P. Monia
TITLE OF INVENTION: ANTISENSE MODULATION OF HUMAN MDN2
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 271
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Jane Massey Licata
STREET: 66 East Main Street
CITY: Marlton
STATE: NJ
COUNTRY: U.S.A.
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PC
OPERATING SYSTEM: WINDOWS 95
SOFTWARE: WORDPERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/752,983
CLASSIFICATION:
FILING DATE: 02-Jan-2001
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/280,805
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Licata, Jane Massey
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0346
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-810-1515
TELEFAX: 609-810-1454
INFORMATION FOR SEQ ID NO: 266:
SEQUENCE CHARACTERISTICS:

LENGTH: 20 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: Yes
US-09-752-983-266

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TCCGACACTTGGGAGGCC 1877
DB 1 TCCGACACTTGGGAGGCC 20

RESULT 2
US-09-752-983-267
Sequence 267, Application US/09752983
Patent No. US20010016575A1
GENERAL INFORMATION:

APPLICANT: Loren J. Miraglia, Pamela Nero, Mark J.
APPLICANT: Graham, Brett P. Monia
TITLE OF INVENTION: ANTISENSE MODULATION OF HUMAN MDW2
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 271
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Jane Massey Licata
STREET: 66 East Main Street
CITY: Marlton
STATE: NJ

COUNTRY: U.S.A.

ZIP: 08053

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE

COMPUTER: IBM PC

OPERATING SYSTEM: WINDOWS 95

SOFTWARE: WORDPERFECT 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/752,983

FILING DATE: 02-Jan-2001

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/280,805

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Licata, Jane Massey

REGISTRATION NUMBER: 32,257

REFERENCE/DOCKET NUMBER: ISPH-0346

TELECOMMUNICATION INFORMATION:

TELEPHONE: 609-810-1515

TELEFAX: 609-810-1454

INFORMATION FOR SEQ ID NO: 267:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

ANTI-SENSE: Yes

US-09-752-983-267

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1851 CCTGTATCCGACACTTGG 1870
DB 1 CCTGTATCCGACACTTGG 20

RESULT 3
US-09-733-294A-82/c

Sequence 82, Application US/09733294A

Patent No. US20020045588A1
GENERAL INFORMATION:
APPLICANT: Brett P. Monia
APPLICANT: William Gaarde
APPLICANT: Susan M. Freier
APPLICANT: Edward V. Mancewicz
TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION
FILE REFERENCE: ISPH-0527
CURRENT APPLICATION NUMBER: US/09/733,294A
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: 09/572,423
PRIOR FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS: 108
SEQ ID NO 82
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antisense Oligonucleotide
US-09-733-294A-82

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1846 TCACGCGCTGTATCCGACA 1865
DB 20 TCACGCGCTGTATCCGACA 1

RESULT 4
US-09-973-788A-55
Sequence 55, Application US/09973788A
Patent No. US20020127574A1
GENERAL INFORMATION:

APPLICANT: Markin, Chad A.
APPLICANT: Letsinger, Robert L.
APPLICANT: Mucic, Robert C.
APPLICANT: Stornoff, James J.
APPLICANT: Elghanian, Robert

APPLICANT: Taton, Thomas A.
TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: 00-713-110
CURRENT APPLICATION NUMBER: US/09/973,788A

CURRENT FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: 09/603,830
PRIOR FILING DATE: 2000-06-26
PRIOR APPLICATION NUMBER: 09/344,667
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 09/240,755
PRIOR FILING DATE: 1999-01-29

PRIOR APPLICATION NUMBER: PCT/US97/12783
PRIOR FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/031,809
PRIOR FILING DATE: 1996-07-29
PRIOR APPLICATION NUMBER: 60/200,161
PRIOR FILING DATE: 2000-04-26
NUMBER OF SEQ ID NOS: 64
SOFTWARE: Microsoft Word 2000
SEQ ID NO 55
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: random
OTHER INFORMATION: synthetic sequence
US-09-973-788A-55

Query Match
Best Local Similarity 100.0%; Score 20; DB 9; Length 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1799 AAAAAAAAAAAAAAAAAA 1818
|||
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 5
US-09-973-638A-55
; Sequence 55, Application US/09973638A
; Patent No. US2002013707A1
; GENERAL INFORMATION:
; APPLICANT: Mirkin, Chad A.
; APPLICANT: Letsinger, Robert L.
; APPLICANT: Mucic, Robert C.
; APPLICANT: Storchoff, James J.
; APPLICANT: Elghanian, Robert
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
; FILE REFERENCE: 00-713-19
; CURRENT APPLICATION NUMBER: US/09/973,638A
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/603,830
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 09/344,667
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/240,755
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: PCT/US97/12783
; PRIOR FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/031,809
; PRIOR FILING DATE: 1996-07-29
; PRIOR APPLICATION NUMBER: 60/200,161
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Microsoft Word 2000
; SEQ ID NO 55
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: random
; OTHER INFORMATION: Synthetic sequence
US-09-973-638A-55

Query Match 0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818
|||
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 6
US-09-974-007-55
; Sequence 55, Application US/09974007
; Patent No. US2002013707A1
; GENERAL INFORMATION:
; APPLICANT: Mirkin, Chad A.
; APPLICANT: Letsinger, Robert L.
; APPLICANT: Mucic, Robert C.
; APPLICANT: Storchoff, James J.
; APPLICANT: Elghanian, Robert
; APPLICANT: Taton, Thomas A.
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
; FILE REFERENCE: 00-713-18
; CURRENT APPLICATION NUMBER: US/09/974,007
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/603,830
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 09/344,667
; PRIOR FILING DATE: 1999-06-25
```

```
; PRIOR APPLICATION NUMBER: 09/240,755
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: PCT/US97/12783
; PRIOR FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/031,809
; PRIOR FILING DATE: 1996-07-29
; PRIOR APPLICATION NUMBER: 60/200,161
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Microsoft Word 2000
; SEQ ID NO 55
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: random
; OTHER INFORMATION: Synthetic sequence
US-09-974-007-55

Query Match 0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818
|||
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 7
US-09-976-617A-55
; Sequence 55, Application US/09976617A
; Patent No. US2002013707A1
; GENERAL INFORMATION:
; APPLICANT: Mirkin, Chad A.
; APPLICANT: Letsinger, Robert L.
; APPLICANT: Mucic, Robert C.
; APPLICANT: Storchoff, James J.
; APPLICANT: Elghanian, Robert
; APPLICANT: Taton, Thomas A.
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
; FILE REFERENCE: 00-713-124
; CURRENT APPLICATION NUMBER: US/09/976,617A
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/603,830
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 09/344,667
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/240,755
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: PCT/US97/12783
; PRIOR FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/031,809
; PRIOR FILING DATE: 1996-07-29
; PRIOR APPLICATION NUMBER: 60/200,161
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Microsoft Word 2000
; SEQ ID NO 55
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: random
; OTHER INFORMATION: Synthetic sequence
US-09-976-617A-55

Query Match 0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

DB 1 AAAAAAAAAAAAAAAAAA 20

RESULT 8
US-09-961-949A-55
Sequence 55, Application US/09961949A
Patent No. US20020146720A1
GENERAL INFORMATION:
APPLICANT: Mirkin, Chad A.
APPLICANT: Letsinger, Robert L.
APPLICANT: Mucic, Robert C.
APPLICANT: Storchoff, James J.
APPLICANT: Elghanian, Robert
APPLICANT: Taton, Thomas A.
TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
FILE REFERENCE: 00-713-11
CURRENT APPLICATION NUMBER: US/09/961,949A
CURRENT FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: 09/603,830
PRIOR FILING DATE: 2000-06-26
PRIOR APPLICATION NUMBER: 09/344,667
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 09/240,755
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: PCT/US97/12783
PRIOR FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/031,809
PRIOR FILING DATE: 1996-07-29
PRIOR APPLICATION NUMBER: 60/200,161
PRIOR FILING DATE: 2000-04-26
NUMBER OF SEQ ID NOS: 64
SOFTWARE: Microsoft Word 2000
SEQ ID NO 55
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: random
US-09-961-949A-55

Query Match 0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
1 AAAAAAAAAAAAAAAAAA 20

DB 1 AAAAAAAAAAAAAAAAAA 20

RESULT 9
US-09-760-500A-55
Sequence 55, Application US/09760500A
Patent No. US2002015542A1
GENERAL INFORMATION:
APPLICANT: Mirkin, Chad A.
APPLICANT: Letsinger, Robert L.
APPLICANT: Mucic, Robert C.
APPLICANT: Storchoff, James J.
APPLICANT: Elghanian, Robert
APPLICANT: Taton, Thomas A.
TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
FILE REFERENCE: 00-715-A
CURRENT APPLICATION NUMBER: US/09/760,500A
CURRENT FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: 09/603,830
PRIOR FILING DATE: 2000-06-26
PRIOR APPLICATION NUMBER: 09/344,667
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 09/240,755
PRIOR FILING DATE: 1999-01-29

PRIOR APPLICATION NUMBER: PCT/US97/12783
PRIOR FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/031,809
PRIOR FILING DATE: 1996-07-29
PRIOR APPLICATION NUMBER: 60/200,161
PRIOR FILING DATE: 2000-04-26
NUMBER OF SEQ ID NOS: 64
SOFTWARE: Microsoft Word 2000
SEQ ID NO 55
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: random
US-09-760-500A-55

Query Match 0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
1 AAAAAAAAAAAAAAAAAA 20

DB 1 AAAAAAAAAAAAAAAAAA 20

RESULT 10
US-09-967-409A-55
Sequence 55, Application US/09967409A
Patent No. US20020155458A1
GENERAL INFORMATION:
APPLICANT: Mirkin, Chad A.
APPLICANT: Letsinger, Robert L.
APPLICANT: Mucic, Robert C.
APPLICANT: Storchoff, James J.
APPLICANT: Elghanian, Robert
APPLICANT: Taton, Thomas A.
TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
FILE REFERENCE: 00-713-16
CURRENT APPLICATION NUMBER: US/09/967,409A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: 09/603,830
PRIOR FILING DATE: 2000-06-26
PRIOR APPLICATION NUMBER: 09/344,667
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 09/240,755
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: PCT/US97/12783
PRIOR FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/031,809
PRIOR FILING DATE: 1996-07-29
PRIOR APPLICATION NUMBER: 60/200,161
PRIOR FILING DATE: 2000-04-26
NUMBER OF SEQ ID NOS: 64
SOFTWARE: Microsoft Word 2000
SEQ ID NO 55
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: random
US-09-967-409A-55

Query Match 0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
1 AAAAAAAAAAAAAAAAAA 20

DB 1 AAAAAAAAAAAAAAAAAA 20

```
RESULT 11
US-09-975-062A-55
; Sequence 55, Application US/09975062A
; Patent No. US20020155459A1
; GENERAL INFORMATION:
; APPLICANT: Mirkin, Chad A.
; APPLICANT: Letsinger, Robert L.
; APPLICANT: Mucic, Robert C.
; APPLICANT: Storchoff, James J.
; APPLICANT: Elghanian, Robert
; APPLICANT: Taton, Thomas A.
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
; FILE REFERENCE: 00-713-111
; CURRENT APPLICATION NUMBER: US/09/975,062A
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/603,830
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 09/344,667
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/240,755
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: PCT/US97/12783
; PRIOR FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/031,809
; PRIOR FILING DATE: 1996-07-29
; PRIOR APPLICATION NUMBER: 60/200,161
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Microsoft Word 2000
; SEQ ID NO 55
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:random
US-09-975-062A-55

Query Match          0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 12
US-09-976-378A-55
; Sequence 55, Application US/09976378A
; Patent No. US20020155461A1
; GENERAL INFORMATION:
; APPLICANT: Mirkin, Chad A.
; APPLICANT: Letsinger, Robert L.
; APPLICANT: Mucic, Robert C.
; APPLICANT: Storchoff, James J.
; APPLICANT: Elghanian, Robert
; APPLICANT: Taton, Thomas A.
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
; FILE REFERENCE: 00-713-125
; CURRENT APPLICATION NUMBER: US/09/976,378A
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/603,830
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 09/344,667
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/240,755
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: PCT/US97/12783
; PRIOR FILING DATE: 1997-07-21
```

```

; PRIOR APPLICATION NUMBER: 60/031,809
; PRIOR FILING DATE: 1996-07-29
; PRIOR APPLICATION NUMBER: 60/200,161
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Microsoft Word 2000
; SEQ ID NO 55
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:random
US-09-976-378A-55

Query Match          0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 13
US-09-976-577-55
; Sequence 55, Application US/09976577
; Patent No. US20020155462A1
; GENERAL INFORMATION:
; APPLICANT: Mirkin, Chad A.
; APPLICANT: Letsinger, Robert L.
; APPLICANT: Mucic, Robert C.
; APPLICANT: Storchoff, James J.
; APPLICANT: Elghanian, Robert
; APPLICANT: Taton, Thomas A.
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
; FILE REFERENCE: 00-713-120
; CURRENT APPLICATION NUMBER: US/09/976,577
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/603,830
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 09/344,667
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/240,755
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: PCT/US97/12783
; PRIOR FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/031,809
; PRIOR FILING DATE: 1996-07-29
; PRIOR APPLICATION NUMBER: 60/200,161
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Microsoft Word 2000
; SEQ ID NO 55
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:random
US-09-976-577-55

Query Match          0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 14
```

```
US-09-771-554-5/c
; Sequence 5, Application US/09771554
; Patent No. US20020155496A1
; GENERAL INFORMATION:
; APPLICANT: CHARLES, Marie Helene
; APPLICANT: FIGA, Nadia
; APPLICANT: BATTAIL-FOIROT, Nicole
; APPLICANT: VERON, Laurent
; APPLICANT: DELAIR, Thierry
; APPLICANT: MANDRAND, Bernard
; TITLE OF INVENTION: SATURATED AND UNSATURATED ABITANE DERIVATIVES, DERIVED CONJUGATE
; FILE REFERENCE: 108473
; CURRENT APPLICATION NUMBER: US/09/771,554
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: PCT/FR99/01846
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: FR 98/10084
; PRIOR FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-09-771-554-5

Query Match          0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 15
US-09-966-312-55
; Sequence 55, Application US/09966312
; Patent No. US20020164605A1
; GENERAL INFORMATION:
; APPLICANT: Mirkin, Chad A.
; APPLICANT: Letsinger, Robert L.
; APPLICANT: Mucic, Robert C.
; APPLICANT: Storchoff, James J.
; APPLICANT: Bishanian, Robert
; APPLICANT: Tator, Thomas A.
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
; FILE REFERENCE: 00-713-15
; CURRENT APPLICATION NUMBER: US/09/966,312
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 09/603,830
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 09/344,667
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/240,755
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: PCT/US97/12783
; PRIOR FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/031,809
; PRIOR FILING DATE: 1996-07-29
; PRIOR APPLICATION NUMBER: 60/200,161
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Microsoft Word 2000
; SEQ ID NO 55
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: Description of Artificial Sequence: random
; OTHER INFORMATION: synthetic sequence
US-09-966-312-55

Query Match          0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

Search completed: March 7, 2004, 23:00:09
Job time : 964 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 14:48:48 ; Search time 9053 Seconds
(without alignments)
10952.666 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818

Sequence: 1 agagaccatcatatctga.....aaaaaaaaaaaaaaaaaaaa 2818

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 37577330 seqs, 17593059518 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2688236

Minimum DB seq length: 0

Maximum DB seq length: 20

Post-Processing: Listing first 45 summaries

Database :

Pending Patents NA Main:*

- 1: /cgn2_6/ptodata/2/pna/US1008_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
- 9: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
- 10: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
- 11: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
- 12: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
- 13: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
- 14: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
- 15: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
- 16: /cgn2_6/ptodata/2/pna/US092_COMB.seq.*
- 17: /cgn2_6/ptodata/2/pna/US093_COMB.seq.*
- 18: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
- 19: /cgn2_6/ptodata/2/pna/US095_COMB.seq.*
- 20: /cgn2_6/ptodata/2/pna/US096_COMB.seq.*
- 21: /cgn2_6/ptodata/2/pna/US097_COMB.seq.*
- 22: /cgn2_6/ptodata/2/pna/US098_COMB.seq.*
- 23: /cgn2_6/ptodata/2/pna/US099_COMB.seq.*
- 24: /cgn2_6/ptodata/2/pna/US100_COMB.seq.*
- 25: /cgn2_6/ptodata/2/pna/US101_COMB.seq.*
- 26: /cgn2_6/ptodata/2/pna/US102_COMB.seq.*
- 27: /cgn2_6/ptodata/2/pna/US103_COMB.seq.*
- 28: /cgn2_6/ptodata/2/pna/US104_COMB.seq.*
- 29: /cgn2_6/ptodata/2/pna/US105_COMB.seq.*
- 30: /cgn2_6/ptodata/2/pna/US106_COMB.seq.*
- 31: /cgn2_6/ptodata/2/pna/US107_COMB.seq.*
- 32: /cgn2_6/ptodata/2/pna/US108_COMB.seq.*
- 33: /cgn2_6/ptodata/2/pna/US109_COMB.seq.*
- 34: /cgn2_6/ptodata/2/pna/US110_COMB.seq.*
- 35: /cgn2_6/ptodata/2/pna/US111_COMB.seq.*
- 36: /cgn2_6/ptodata/2/pna/US112_COMB.seq.*
- 37: /cgn2_6/ptodata/2/pna/US113_COMB.seq.*
- 38: /cgn2_6/ptodata/2/pna/US114_COMB.seq.*
- 39: /cgn2_6/ptodata/2/pna/US115_COMB.seq.*
- 40: /cgn2_6/ptodata/2/pna/US116_COMB.seq.*
- 41: /cgn2_6/ptodata/2/pna/US117_COMB.seq.*
- 42: /cgn2_6/ptodata/2/pna/US118_COMB.seq.*
- 43: /cgn2_6/ptodata/2/pna/US119_COMB.seq.*

- 44: /cgn2_6/ptodata/2/pna/US1008_COMB.seq.*
- 45: /cgn2_6/ptodata/2/pna/US1018_COMB.seq.*
- 46: /cgn2_6/ptodata/2/pna/US1019_COMB.seq.*
- 47: /cgn2_6/ptodata/2/pna/US1028_COMB.seq.*
- 48: /cgn2_6/ptodata/2/pna/US1038_COMB.seq.*
- 49: /cgn2_6/ptodata/2/pna/US1039_COMB.seq.*
- 50: /cgn2_6/ptodata/2/pna/US103B_COMB.seq.*
- 51: /cgn2_6/ptodata/2/pna/US104A_COMB.seq.*
- 52: /cgn2_6/ptodata/2/pna/US104B_COMB.seq.*
- 53: /cgn2_6/ptodata/2/pna/US106_COMB.seq.*
- 54: /cgn2_6/ptodata/2/pna/US107A_COMB.seq.*
- 55: /cgn2_6/ptodata/2/pna/US107B_COMB.seq.*
- 56: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
- 57: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
- 58: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
- 59: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
- 60: /cgn2_6/ptodata/2/pna/US6004_COMB.seq.*
- 61: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
- 62: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
- 63: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
- 64: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
- 65: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
- 66: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
- 67: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
- 68: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
- 69: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
- 70: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
- 71: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
- 72: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
- 73: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
- 74: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
- 75: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
- 76: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
- 77: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
- 78: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
- 79: /cgn2_6/ptodata/2/pna/US6023_COMB.seq.*
- 80: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
- 81: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
- 82: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
- 83: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*
- 84: /cgn2_6/ptodata/2/pna/US6037_COMB.seq.*
- 85: /cgn2_6/ptodata/2/pna/US6038_COMB.seq.*
- 86: /cgn2_6/ptodata/2/pna/US6039_COMB.seq.*
- 87: /cgn2_6/ptodata/2/pna/US6041_COMB.seq.*
- 88: /cgn2_6/ptodata/2/pna/US6042_COMB.seq.*
- 89: /cgn2_6/ptodata/2/pna/US6043_COMB.seq.*
- 90: /cgn2_6/ptodata/2/pna/US6044_COMB.seq.*
- 91: /cgn2_6/ptodata/2/pna/US6045_COMB.seq.*
- 92: /cgn2_6/ptodata/2/pna/US6046_COMB.seq.*
- 93: /cgn2_6/ptodata/2/pna/US6047_COMB.seq.*
- 94: /cgn2_6/ptodata/2/pna/US6048_COMB.seq.*
- 95: /cgn2_6/ptodata/2/pna/US6049_COMB.seq.*
- 96: /cgn2_6/ptodata/2/pna/US6050_COMB.seq.*
- 97: /cgn2_6/ptodata/2/pna/US6051_COMB.seq.*
- 98: /cgn2_6/ptodata/2/pna/US6052_COMB.seq.*
- 99: /cgn2_6/ptodata/2/pna/US6053_COMB.seq.*
- 100: /cgn2_6/ptodata/2/pna/US6054_COMB.seq.*
- 101: /cgn2_6/ptodata/2/pna/US6055_COMB.seq.*
- 102: /cgn2_6/ptodata/2/pna/US6056_COMB.seq.*
- 103: /cgn2_6/ptodata/2/pna/US6057_COMB.seq.*
- 104: /cgn2_6/ptodata/2/pna/US6058_COMB.seq.*
- 105: /cgn2_6/ptodata/2/pna/US6059_COMB.seq.*
- 106: /cgn2_6/ptodata/2/pna/US6060_COMB.seq.*
- 107: /cgn2_6/ptodata/2/pna/US6061_COMB.seq.*
- 108: /cgn2_6/ptodata/2/pna/US6062_COMB.seq.*
- 109: /cgn2_6/ptodata/2/pna/US6063_COMB.seq.*
- 110: /cgn2_6/ptodata/2/pna/US6064_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	0.7	20	1	PCT-US00-07938-10	Sequence 10, Appl
2	20	0.7	20	1	PCT-US00-17507-53	Sequence 53, Appl
3	20	0.7	20	1	PCT-US01-01411A-62	Sequence 62, Appl
4	20	0.7	20	1	PCT-US01-15774-82	Sequence 82, Appl
5	20	0.7	20	1	PCT-US01-30871-70	Sequence 70, Appl
6	20	0.7	20	1	PCT-US01-32233A-23	Sequence 23, Appl
7	20	0.7	20	1	PCT-US01-44737-36	Sequence 36, Appl
8	20	0.7	20	1	PCT-US02-10873A-574	Sequence 574, Appl
9	20	0.7	20	1	PCT-US02-18229-5	Sequence 5, Appl
10	20	0.7	20	1	PCT-US02-18229-6	Sequence 6, Appl
11	20	0.7	20	1	PCT-US02-33002-2	Sequence 2, Appl
12	20	0.7	20	1	PCT-US02-33699-1	Sequence 1, Appl
13	20	0.7	20	1	PCT-US02-35719-13	Sequence 13, Appl
14	20	0.7	20	1	PCT-US02-38281-265	Sequence 265, Appl
15	20	0.7	20	1	PCT-US02-38281-267	Sequence 267, Appl
16	20	0.7	20	1	PCT-US02-39280-36	Sequence 36, Appl
17	20	0.7	20	1	PCT-US03-14100A-31	Sequence 31, Appl
18	20	0.7	20	1	PCT-US03-14100B-31	Sequence 31, Appl
19	20	0.7	20	1	PCT-US03-19360-87	Sequence 87, Appl
20	20	0.7	20	1	PCT-US03-19360-88	Sequence 88, Appl
21	20	0.7	20	1	PCT-US03-19360-222	Sequence 222, Appl
22	20	0.7	20	1	PCT-US03-19360-223	Sequence 223, Appl
23	20	0.7	20	1	PCT-US03-26491-516	Sequence 516, Appl
24	20	0.7	20	1	PCT-US03-27520-1	Sequence 1, Appl
25	20	0.7	20	1	PCT-US03-30374-178	Sequence 178, Appl
26	20	0.7	20	1	PCT-US03-30374-179	Sequence 179, Appl
27	20	0.7	20	1	PCT-US03-30374-180	Sequence 180, Appl
28	20	0.7	20	1	PCT-US03-30374-181	Sequence 181, Appl
29	20	0.7	20	1	PCT-US03-30374-182	Sequence 182, Appl
30	20	0.7	20	1	PCT-US03-30374-183	Sequence 183, Appl
31	20	0.7	20	1	PCT-US03-30374-184	Sequence 184, Appl
32	20	0.7	20	1	PCT-US03-30374-185	Sequence 185, Appl
33	20	0.7	20	1	PCT-US03-30374-186	Sequence 186, Appl
34	20	0.7	20	1	PCT-US03-30374-187	Sequence 187, Appl
35	20	0.7	20	1	PCT-US03-30374-188	Sequence 188, Appl
36	20	0.7	20	1	PCT-US03-30374-189	Sequence 189, Appl
37	20	0.7	20	1	PCT-US03-30374-190	Sequence 190, Appl
38	20	0.7	20	1	PCT-US03-30374-191	Sequence 191, Appl
39	20	0.7	20	1	PCT-US03-30374-192	Sequence 192, Appl
40	20	0.7	20	1	PCT-US03-30374-193	Sequence 193, Appl
41	20	0.7	20	1	PCT-US03-30374-194	Sequence 194, Appl
42	20	0.7	20	1	PCT-US03-30374-195	Sequence 195, Appl
43	20	0.7	20	1	PCT-US03-30374-196	Sequence 196, Appl
44	20	0.7	20	1	PCT-US03-30374-197	Sequence 197, Appl
45	20	0.7	20	1	PCT-US03-30374-198	Sequence 198, Appl

ALIGNMENTS

RESULT 1
PCT-US00-07938-10
; Sequence 10, Application PC/TUS0007938
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: MICROSATELLITE REPEATS IN THE COSTIMULATORY RECEPTOR LOCUS AND U
; FILE REFERENCE: GNN-5343CPC
; CURRENT FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/1126,215
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: primer
PCT-US00-07938-10

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2072 GTTGCACTCCAGCTGGCG 2091
Db 1 GTTGCACTCCAGCTGGCG 20

RESULT 2
PCT-US00-17507-53
; Sequence 53, Application PC/TUS0017507
; GENERAL INFORMATION:
; APPLICANT: Markin, Chad A.
; APPLICANT: Letsinger, Robert L.
; APPLICANT: Mucic, Robert C.
; APPLICANT: Storch, James J.
; APPLICANT: Elshanian, Robert
; APPLICANT: Talon, Thomas A.
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
; FILE REFERENCE: 4149-1-1-1-PCT
; CURRENT FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 09/344,667
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/200,161
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: random
PCT-US00-17507-53

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 3
PCT-US01-01411A-62/c
; Sequence 62, Application PC/TUS0101411A
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Brett P. Monia
; APPLICANT: Robert McKay
; APPLICANT: Madeline W. Butler
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF GLYCOGEN SYNTHASE KINASE 3 ALPHA
; FILE REFERENCE: RTSP-0096
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 09/489,856
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 62
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

OTHER INFORMATION: Antisense Oligonucleotide
PCT-US01-01411A-62

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1854 GTATCCCGACACTTGGG 1873
DB 20 GTATCCCGACACTTGGG 1

RESULT 4
PCT-US01-15774-82/c
Sequence 82, Application PC/TUS0115774
GENERAL INFORMATION:
APPLICANT: Isis Pharmaceuticals, Inc.
APPLICANT: Brett P. Monia
APPLICANT: William Gaarde
APPLICANT: Susan M. Freiler
APPLICANT: Edward V. Wanciewicz
TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION
FILE REFERENCE: ISPH-0568
CURRENT APPLICATION NUMBER: PCT/US01/15774
CURRENT FILING DATE: 2001-05-15
PRIOR FILING DATE: 09/572,423
PRIOR FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS: 108
SEQ ID NO 82
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antisense Oligonucleotide
PCT-US01-15774-82

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1846 TCACGCGTGTATCCGACA 1865
DB 20 TCACGCGTGTATCCGACA 1

RESULT 5
PCT-US01-30871-70/c
Sequence 70, Application PC/TUS0130871
GENERAL INFORMATION:
APPLICANT: Isis Pharmaceuticals, Inc.
APPLICANT: Vickie L. Brown-Driver
APPLICANT: Hong Zhang
APPLICANT: Andrew T. Watt
TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 6 EXPRESSION
FILE REFERENCE: RISP-0192
CURRENT APPLICATION NUMBER: PCT/US01/30871
CURRENT FILING DATE: 2001-10-03
PRIOR APPLICATION NUMBER: 09/679,299
PRIOR FILING DATE: 2000-10-04
NUMBER OF SEQ ID NOS: 164
SEQ ID NO 70
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antisense Oligonucleotide
PCT-US01-30871-70

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1852 CTGTATCCCGACACTTGG 1871

DB 20 CTGTATCCCGACACTTGG 1

RESULT 6
PCT-US01-32233A-23/c
Sequence 23, Application PC/TUS0132233A
GENERAL INFORMATION:
APPLICANT: Phyllos, Inc.
TITLE OF INVENTION: PROTEIN SCAFFOLDS FOR ANTIBODY MIMICS
TITLE OF INVENTION: AND OTHER BINDING PROTEINS
FILE REFERENCE: 50036/021M04
CURRENT APPLICATION NUMBER: PCT/US01/32233A
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: US 09/688,566
PRIOR FILING DATE: 2000-10-16
NUMBER OF SEQ ID NOS: 202
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 20
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US01-32233A-23

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1799 AAAAAAAAAAAAAAAAAA 1818
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 7
PCT-US01-44737-36/c
Sequence 36, Application PC/TUS0144737
GENERAL INFORMATION:
APPLICANT: Weterings, Koen
APPLICANT: Apuya, Nestor R.
APPLICANT: Tatariyova, Tatiana
APPLICANT: Goldberg, Robert B.
APPLICANT: The Regents of the University of California
APPLICANT: Ceres, Inc.
TITLE OF INVENTION: Polynucleotides Useful for Modulating Transcription
FILE REFERENCE: 023070-114710PC
CURRENT APPLICATION NUMBER: PCT/US01/44737
CURRENT FILING DATE: 2002-08-28
PRIOR APPLICATION NUMBER: US 09/724,857
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: US 60/253,672
PRIOR FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 36
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: oligo (dT-20)
PCT-US01-44737-36

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1799 AAAAAAAAAAAAAAAAAA 1818
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 8
PCT-US02-10873A-574

```
; Sequence 574, Application PC/TUS0210873A
; GENERAL INFORMATION:
; APPLICANT: Byrne, David J.
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Evans, Claire F.
; APPLICANT: Mah, Steven P.
; APPLICANT: Lo, David D.
; TITLE OF INVENTION: Genes Expressed in Intestinal Epithelium and Peyer's Patch M Cell
; FILE REFERENCE: 216019-109
; CURRENT APPLICATION NUMBER: PCT/US02/10873A
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/281,416
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 859
; SOFTWARE: PatentIn Version 3.1
; SEQ ID NO 574
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Real-time validation FOR primer for EDDIS_87
PCT-US02-10873A-574
```

```
Query Match          0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred.No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

QY      1845 CTCACGCGCTGTAATCCAGC 1864
DB      1 CTCACGCGCTGTAATCCAGC 20
```

```
RESULT 9
PCT-US02-18229-5
; Sequence 5, Application PC/TUS0218229
; GENERAL INFORMATION:
; APPLICANT: Haydock, Paul V.
; APPLICANT: U'Ren, Jack
; APPLICANT: Salsigne Corporation
; TITLE OF INVENTION: Nucleic Acid Amplification Utilizing Intermediate
; FILE REFERENCE: 018048-001711PC
; CURRENT APPLICATION NUMBER: PCT/US02/18229
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/296,812
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 10/077,383
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: (A)-12-20
; OTHER INFORMATION: homopolymer spacer sequence
; NAME/KEY: modified base
; LOCATION: (13)..(20)
; OTHER INFORMATION: a at positions 13-20 may be present or absent
PCT-US02-18229-5
```

```
Query Match          0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred.No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

QY      1799 AAAAAAAAAAAAAAAAAAAAAA 1818
DB      1 AAAAAAAAAAAAAAAAAAAAAA 20
```

```
RESULT 10
PCT-US02-18229-6/c
; Sequence 6, Application PC/TUS0218229
; GENERAL INFORMATION:
; APPLICANT: Haydock, Paul V.
; APPLICANT: U'Ren, Jack
; APPLICANT: Salsigne Corporation
; TITLE OF INVENTION: Nucleic Acid Amplification Utilizing Intermediate
; FILE REFERENCE: 018048-001711PC
; CURRENT APPLICATION NUMBER: PCT/US02/18229
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/296,812
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 10/077,383
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: (T)-12-20
; NAME/KEY: modified base
; LOCATION: (13)..(20)
; OTHER INFORMATION: t at positions 13-20 may be present or absent
PCT-US02-18229-6
```

```
Query Match          0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred.No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

QY      1799 AAAAAAAAAAAAAAAAAAAAAA 1818
DB      20 AAAAAAAAAAAAAAAAAAAAAA 1
```

```
RESULT 11
PCT-US02-33002-2/c
; Sequence 2, Application PC/TUS0233002
; GENERAL INFORMATION:
; APPLICANT: Serafini, Tito
; APPLICANT: Chen, Hang
; APPLICANT: Toemel, Emily
; APPLICANT: Ellis, Michael
; TITLE OF INVENTION: Methods for Nucleic Acid Amplification
; FILE REFERENCE: 2605-1-002PCT
; CURRENT APPLICATION NUMBER: PCT/US02/33002
; CURRENT FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: US 10/036,860
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/305,666
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic sequence
PCT-US02-33002-2
```

```
Query Match          0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred.No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

QY      1799 AAAAAAAAAAAAAAAAAAAAAA 1818
DB      20 AAAAAAAAAAAAAAAAAAAAAA 1
```

RESULT 12
PCT-US02-33699-1/c
; Sequence 1, Application PC/TUS0233699
; GENERAL INFORMATION:
; APPLICANT: Prologo, LLC
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHODS TO DETECT AND/OR QUANTIFY NUCLEIC
; FILE REFERENCE: ACID ANALYTES
; CURRENT APPLICATION NUMBER: PCT/US02/33699
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: 60/336,432
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent version 3.2
; SEQ ID NO 1
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Nucleic Acid Probe
; NAME/KEY: misc.feature
; LOCATION: (1)..(20)
PCT-US02-33699-1

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 13
PCT-US02-35719-13
; Sequence 13, Application PC/TUS0235719
; GENERAL INFORMATION:
; APPLICANT: Wood, Linda
; APPLICANT: Wagner, Susanne
; APPLICANT: Parodi, Luis
; TITLE OF INVENTION: Single Nucleotide Polymorphisms in GH-1
; FILE REFERENCE: 00791.US1
; CURRENT APPLICATION NUMBER: PCT/US02/35719
; CURRENT FILING DATE: 2002-11-07
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent version 3.1
; SEQ ID NO 13
; LENGTH: 20
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: primer
PCT-US02-35719-13

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1852 CTGTAATCCAGCACTTTGG 1871
DB 1 CTGTAATCCAGCACTTTGG 20

RESULT 14
PCT-US02-38281-266
; Sequence 266, Application PC/TUS0238281
; GENERAL INFORMATION:
; APPLICANT: Loren J. Miraglia
; APPLICANT: Pamela S. Nero
; APPLICANT: Mark J. Graham
PCT-US02-38281-266

APPLICANT: Brett P. Monia
APPLICANT: Rich Koller
APPLICANT: Ming Yi Chiang
APPLICANT: Muthiah Manoharan
APPLICANT: Isis Pharmaceuticals, Inc.
TITLE OF INVENTION: Antisense Modulation of Human mdm2 Expression
FILE REFERENCE: ISPH-0717
CURRENT APPLICATION NUMBER: PCT/US02/38281
CURRENT FILING DATE: 2002-12-02
PRIOR APPLICATION NUMBER: US 10/005,344
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: US 09/752,983
PRIOR FILING DATE: 2001-01-02
PRIOR APPLICATION NUMBER: US 09/280,805
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: US 09/048,810
PRIOR FILING DATE: 1998-03-26
NUMBER OF SEQ ID NOS: 379
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 266
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antisense Oligonucleotide
PCT-US02-38281-266

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TCCGAGCACTTTGGAGGCC 1877
DB 1 TCCGAGCACTTTGGAGGCC 20

RESULT 15
PCT-US02-38281-267
; Sequence 267, Application PC/TUS0238281
; GENERAL INFORMATION:
; APPLICANT: Loren J. Miraglia
; APPLICANT: Pamela S. Nero
; APPLICANT: Mark J. Graham
; APPLICANT: Brett P. Monia
; APPLICANT: Rich Koller
; APPLICANT: Ming Yi Chiang
; APPLICANT: Muthiah Manoharan
; APPLICANT: Isis Pharmaceuticals, Inc.
TITLE OF INVENTION: Antisense Modulation of Human mdm2 Expression
FILE REFERENCE: ISPH-0717
CURRENT APPLICATION NUMBER: PCT/US02/38281
CURRENT FILING DATE: 2002-12-02
PRIOR APPLICATION NUMBER: US 10/005,344
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: US 09/752,983
PRIOR FILING DATE: 2001-01-02
PRIOR APPLICATION NUMBER: US 09/280,805
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: US 09/048,810
PRIOR FILING DATE: 1998-03-26
NUMBER OF SEQ ID NOS: 379
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 267
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antisense Oligonucleotide
PCT-US02-38281-267

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1851 CCTGTAATCCAGCACTTG 1870
Db 1 CCTGTAATCCAGCACTTG 20

Search completed: March 7, 2004, 22:42:14
Job time : 9054 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 15:07:23 ; Search time 94 Seconds

(without alignments)
6400.467 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818
Sequence: 1 agagaaccatcatcatgta.....aaaaaaaaaaaaaaaaaa 2818Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 211252 seqs, 106750167 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27452

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Listing first 45 summaries

Database : Pending Patents_NA.New:*

1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	0.7	20	1	PCT-US04-04452-51
2	20	0.7	20	1	PCT-US04-04452-1813
3	20	0.7	20	6	US-10-653-416-25
4	20	0.7	20	6	US-10-213-257-23
5	20	0.7	20	6	US-10-708-204-5116
6	19	0.7	20	5	US-09-574-779B-30
7	19	0.7	20	6	US-10-708-204-5092
8	18	0.6	18	6	US-10-653-416-26
9	18	0.6	18	6	US-10-708-204-3825
10	18	0.6	18	6	US-10-708-204-5130
11	18	0.6	19	1	PCT-US03-25997A-50
12	18	0.6	19	1	PCT-US04-04452-252
13	18	0.6	20	6	US-10-708-204-2896
14	18	0.6	20	6	US-10-708-204-3868
15	17	0.6	19	6	US-10-708-204-3854
16	17	0.6	19	6	US-10-708-204-5172
17	17	0.6	20	6	US-10-708-204-5094
18	16	0.6	16	6	US-10-776-099-9
19	16	0.6	18	6	US-10-708-204-5101
20	16	0.6	19	1	PCT-US04-04452-253
21	16	0.6	19	6	US-10-708-204-3840
22	16	0.6	19	6	US-10-708-204-5110
23	16	0.6	20	6	US-10-708-204-3841
24	16	0.6	20	6	US-10-708-204-3842
25	15	0.5	19	6	US-10-708-204-3842
26	15	0.5	19	6	US-10-708-204-4379

27	15	0.5	20	1	PCT-US04-04452-117	Sequence 117, App
28	15	0.5	20	1	PCT-US04-04452-516	Sequence 516, App
29	15	0.5	20	1	PCT-US04-04452-642	Sequence 642, App
30	15	0.5	20	6	US-10-708-204-3210	Sequence 3210, App
31	15	0.5	20	6	US-10-708-204-4471	Sequence 4471, App
32	15	0.5	20	6	US-10-708-204-5144	Sequence 5144, App
33	14	0.5	14	6	US-10-141-535-2	Sequence 2, App1
34	14	0.5	19	1	PCT-US04-04452-460	Sequence 460, App
35	14	0.5	19	1	PCT-US04-04452-1367	Sequence 1367, App
36	14	0.5	20	6	US-10-769-579-2	Sequence 2, App1
37	13	0.5	17	1	PCT-US03-31862-103	Sequence 103, App
38	13	0.5	17	1	PCT-US03-31862-104	Sequence 104, App
39	13	0.5	18	1	PCT-US03-32805-22	Sequence 22, App1
40	13	0.5	19	1	PCT-US03-02038-164	Sequence 164, App
41	13	0.5	19	6	US-10-708-204-3140	Sequence 3140, App
42	13	0.5	20	1	PCT-US04-02344-110	Sequence 110, App
43	13	0.5	20	1	PCT-US04-04452-1667	Sequence 1667, App
44	13	0.5	20	5	US-09-574-779B-15	Sequence 15, App1
45	13	0.5	20	6	US-10-484-805-9	Sequence 9, App1

ALIGNMENTS

RESULT 1
PCT-US04-04452-51/c
Sequence 51, Application PC/TUS0404452
GENERAL INFORMATION:
APPLICANT: Bardelli, Alberto
APPLICANT: Parsons, Will
APPLICANT: Velculescu, Victor
APPLICANT: Kinzler, Kenneth W.
APPLICANT: Vogelstein, Bert
TITLE OF INVENTION: TYROSINE KINASES IMPLICATED IN CANCERS
FILE REFERENCE: 00107.00327
CURRENT APPLICATION NUMBER: PCT/US04/04452
CURRENT FILING DATE: 2004-02-18
NUMBER OF SEQ ID NOS: 2191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 51
LENGTH: 20
TYPE: DNA
ORGANISM: Homo sapiens
PCT-US04-04452-51

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1901 CAGGAGTTGAGACGAGCT 1920
Db 20 CAGGAGTTGAGACGAGCT 1

RESULT 2
PCT-US04-04452-1813/c
Sequence 1813, Application PC/TUS0404452
GENERAL INFORMATION:
APPLICANT: Bardelli, Alberto
APPLICANT: Parsons, Will
APPLICANT: Velculescu, Victor
APPLICANT: Kinzler, Kenneth W.
APPLICANT: Vogelstein, Bert
TITLE OF INVENTION: TYROSINE KINASES IMPLICATED IN CANCERS
FILE REFERENCE: 00107.00327
CURRENT APPLICATION NUMBER: PCT/US04/04452
CURRENT FILING DATE: 2004-02-18
NUMBER OF SEQ ID NOS: 2191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1813
LENGTH: 20
TYPE: DNA
ORGANISM: Homo sapiens

PCT-US04-04452-1813

Query Match 0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1854 GTATCCGACGACTTTGGGA 1873
DB 20 GTATCCGACGACTTTGGGA 1

RESULT 3

US-10-653-416-25/c
Sequence 25, Application US/10653416
GENERAL INFORMATION:
APPLICANT: RASHTCHIAN, AYOUB
APPLICANT: SCHUSTER, DAVID M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CDNA SYNTHESIS
FILE REFERENCE: 38266-0011
CURRENT APPLICATION NUMBER: US/10/653,416
CURRENT FILING DATE: 2003-09-03
PRIOR APPLICATION NUMBER: 60/407,248
PRIOR FILING DATE: 2002-09-03
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 25
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-653-416-25

Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 4

US-10-213-257-23/c
Sequence 23, Application US/10213257
GENERAL INFORMATION:
APPLICANT: WANG, YUAN-YUAN
APPLICANT: KHOO, KAY-HOUI
APPLICANT: CHEN, SHUI-TEIN
APPLICANT: LIN, CHUN-CHENG
APPLICANT: WONG, CHI-HUEY
APPLICANT: LIN, CHUN-HUNG
APPLICANT: CHEN, HONG-SEN
APPLICANT: TSAI, YOW-FU
TITLE OF INVENTION: IMMUNO-MODULATING ANTITUMOR ACTIVITIES OF GANODERMA
TITLE OF INVENTION: LUCIDIM (REISHI) POLYSACCHARIDES
FILE REFERENCE: 4910-12
CURRENT APPLICATION NUMBER: US/10/213,257
CURRENT FILING DATE: 2002-08-06
PRIOR APPLICATION NUMBER: 60/310,285
PRIOR FILING DATE: 2001-08-06
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 23
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-213-257-23

Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 5

US-10-708-204-5116/c
Sequence 5116, Application US/10708204
GENERAL INFORMATION:
APPLICANT: ROSETTA GENOMICS LTD
TITLE OF INVENTION: BIOINFORMATIALLY DETECTABLE GROUP OF NOVEL REGULATORY
TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
THEREOF
FILE REFERENCE: 55033
CURRENT APPLICATION NUMBER: US/10/708,204
CURRENT FILING DATE: 2004-02-16
NUMBER OF SEQ ID NOS: 7351
SOFTWARE: PatentIn version 3.2
SEQ ID NO 5116
LENGTH: 20
TYPE: RNA
ORGANISM: Homo Sapiens
US-10-708-204-5116

Query Match 0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2068 TGCCGTTGCACTCCAGCTG 2087
DB 20 TGCCGTTGCACTCCAGCTG 1

RESULT 6

US-09-574-779B-30/c
Sequence 30, Application US/09574779B
GENERAL INFORMATION:
APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOL
TITLE OF INVENTION: Novel cDNAs encoding catenin-binding proteins with
TITLE OF INVENTION: Function in signalling and/or gene regulation
FILE REFERENCE: 2676-4415US
CURRENT APPLICATION NUMBER: US/09/574,779B
CURRENT FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: 99201543.8
PRIOR FILING DATE: 1999-05-17
NUMBER OF SEQ ID NOS: 158
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 30
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: primer FVR510F
US-09-574-779B-30

Query Match 0.7%; Score 19; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1846 TCAAGCCTGTAATCCAGC 1864
DB 19 TCAAGCCTGTAATCCAGC 1

RESULT 7

US-10-708-204-5092/c
Sequence 5092, Application US/10708204
GENERAL INFORMATION:
APPLICANT: ROSETTA GENOMICS LTD
US-10-708-204-5092

```

; TITLE OF INVENTION: BIOINFORMATICALLY DETECTABLE GROUP OF NOVEL REGULATORY
; TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
; FILE REFERENCE: 55033
; CURRENT APPLICATION NUMBER: US/10/708,204
; CURRENT FILING DATE: 2004-02-16
; NUMBER OF SEQ ID NOS: 7351
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 5092
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Homo Sapiens
US-10-708-204-5092

```

```

Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 20;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2072 GTTGACCTGCGGCGG 2090
DB 20 GTTGACCTGCGGCGG 2

```

```

RESULT 8
US-10-653-416-26/c
; Sequence 26, Application US/10653416
; GENERAL INFORMATION:
; APPLICANT: RASHTCHIAN, AYDUB
; APPLICANT: SCHUSTER, DAVID M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CDNA SYNTHESIS
; FILE REFERENCE: 38266-0011
; CURRENT APPLICATION NUMBER: US/10/653,416
; CURRENT FILING DATE: 2003-09-03
; PRIOR APPLICATION NUMBER: 60/407,248
; PRIOR FILING DATE: 2002-09-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 26
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: oligonucleotide
; FEATURE:
; OTHER INFORMATION: this sequence may encompass 12-18 nucleotides according
; OTHER INFORMATION: to the specification as filed
US-10-653-416-26

```

```

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1799 AAAAAAAAAAAAAAAAAA 1816
DB 18 AAAAAAAAAAAAAAAAAA 1

```

```

RESULT 9
US-10-708-204-3825/c
; Sequence 3825, Application US/10708204
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; TITLE OF INVENTION: BIOINFORMATICALLY DETECTABLE GROUP OF NOVEL REGULATORY
; TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
; FILE REFERENCE: 55033
; CURRENT APPLICATION NUMBER: US/10/708,204
; CURRENT FILING DATE: 2004-02-16
; NUMBER OF SEQ ID NOS: 7351
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3825
; LENGTH: 18

```

```

; TYPE: RNA
; ORGANISM: Homo Sapiens
US-10-708-204-3825

```

```

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2075 GCACTCCAGCTGGCGGA 2092
DB 18 GCACTCCAGCTGGCGGA 1

```

```

RESULT 10
US-10-708-204-5130/c
; Sequence 5130, Application US/10708204
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; TITLE OF INVENTION: BIOINFORMATICALLY DETECTABLE GROUP OF NOVEL REGULATORY
; TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
; FILE REFERENCE: 55033
; CURRENT APPLICATION NUMBER: US/10/708,204
; CURRENT FILING DATE: 2004-02-16
; NUMBER OF SEQ ID NOS: 7351
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5130
; LENGTH: 18
; TYPE: RNA
; ORGANISM: Homo Sapiens
US-10-708-204-5130

```

```

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1912 GACCAGCTGGCCACAT 1929
DB 18 GACCAGCTGGCCACAT 1

```

```

RESULT 11
PCT-US03-25997A-50/c
; Sequence 50, Application PC/TUS0325997A
; GENERAL INFORMATION:
; APPLICANT: CERES, INC.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY USEFUL FOR
; TITLE OF INVENTION: INCREASING PLANT SIZE AND INCREASING THE NUMBER AND SIZE OF LEA
; FILE REFERENCE: 2750-1573F(PC)
; CURRENT APPLICATION NUMBER: PCT/US03/25997A
; CURRENT FILING DATE: 2003-08-18
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligo primer used in the generation of labeled probes for
; OTHER INFORMATION: hybridization from first-strand cDNA
PCT-US03-25997A-50

```

```

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 19;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1799 AAAAAAAAAAAAAAAAAA 1816
DB 18 AAAAAAAAAAAAAAAAAA 1

```

```

RESULT 12
PCT-US04-04452-252

```

```
; Sequence 252, Application PC/TUS0404452
; GENERAL INFORMATION:
; APPLICANT: Bardelli, Alberto
; APPLICANT: Parsons, Will
; APPLICANT: Velculescu, Victor
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: TYROSINE KINASES IMPLICATED IN CANCERS
; FILE REFERENCE: 001107.00327
; CURRENT APPLICATION NUMBER: PCT/US04/04452
; CURRENT FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 2191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-04452-252
```

```
Query Match          0.6%; Score 18; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2008 AGGCTGAGGAGGAGGAAAT 2025
      |||||
DB      1 AGGCTGAGGAGGAGGAAAT 18
```

```
RESULT 13
US-10-708-204-2896
; Sequence 2896, Application US/10708204
; GENERAL INFORMATION:
```

```
; APPLICANT: ROSETTA GENOMICS LTD
; TITLE OF INVENTION: BIOINFORMATICALLY DETECTABLE GROUP OF NOVEL REGULATORY
; TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
; FILE REFERENCE: 55033
; CURRENT APPLICATION NUMBER: US/10/708,204
; CURRENT FILING DATE: 2004-02-16
; NUMBER OF SEQ ID NOS: 7351
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2896
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Homo Sapiens
US-10-708-204-2896
```

```
Query Match          0.6%; Score 18; DB 6; Length 20;
Best Local Similarity 72.2%; Pred. No. 3.1e+02;
Matches 13; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1851 CCTGTATCCGACACTT 1868
      ||:|||||:|
DB      1 CCUGUAUCCGACACTU 18
```

```
RESULT 14
US-10-708-204-3868/C
; Sequence 3868, Application US/10708204
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; TITLE OF INVENTION: BIOINFORMATICALLY DETECTABLE GROUP OF NOVEL REGULATORY
; TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
; FILE REFERENCE: 55033
; CURRENT APPLICATION NUMBER: US/10/708,204
; CURRENT FILING DATE: 2004-02-16
; NUMBER OF SEQ ID NOS: 7351
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3868
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Homo Sapiens
```

```
US-10-708-204-3868
```

```
Query Match          0.6%; Score 18; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2074 TGCACTCCAGCTGGGGC 2091
      |||||
DB      18 TGCACTCCAGCTGGGGC 1
```

```
RESULT 15
US-10-708-204-3854/C
```

```
; Sequence 3854, Application US/10708204
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; TITLE OF INVENTION: BIOINFORMATICALLY DETECTABLE GROUP OF NOVEL REGULATORY
; TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
; FILE REFERENCE: 55033
; CURRENT APPLICATION NUMBER: US/10/708,204
; CURRENT FILING DATE: 2004-02-16
; NUMBER OF SEQ ID NOS: 7351
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3854
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo Sapiens
US-10-708-204-3854
```

```
Query Match          0.6%; Score 17; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2074 TGCACTCCAGCTGGGGC 2090
      |||||
DB      17 TGCACTCCAGCTGGGGC 1
```

```
Search completed: March 7, 2004, 22:43:56
Job time : 94 secs
```



```

CC sequencing consortium of the German Genome Project
CC s1 sequence also available
CC This clone is available at the RZPD in Berlin
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
CC Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de
XX
FH Key Location/Qualifiers
FH
FT source 1..20
FT /db_xref="taxon:9606"
FT /mol_type="mRNA"
FT /organism="Homo sapiens"
FT /clone_lib="566 (synonym: hEKd2). Vector pAMP1; host
FT X1-2b1ue; sites NotI + SalI"
FT /dev_stage="fetal"
FT /tissue_type="kidney"
XX
SQ Sequence 20 BP; 18 A; 1 C; 0 G; 1 T; 0 other;
XX
Query Match 0.7%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1797 TCAAAAAAAAAAAAAAAAAA 1816
Db 1 TCAAAAAAAAAAAAAAAAAA 20
XX
RESULT 2
ID HSM003226 standard; mRNA; EST; 20 BP.
XX
XX AL038750;
XX
XX AL038750.1
XX
XX 12-MAR-1999 (Rel. 59, Created)
XX 12-MAR-1999 (Rel. 59, Last updated, Version 1)
XX
XX Homo sapiens mRNA; EST DKFZp566M146_r1 (from clone DKFZp566M146)
XX
XX EST; expressed sequence tag.
XX
XX Homo sapiens (human)
XX Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
XX Eutheria; Primates; Catarrhini; Homiidae; Homo.
XX
XX [1]
XX 1-20
XX Otterwaelder B., Obermaier B., Mewes W., Gaassenhuber J., Wiemann S.;
XX Submitted (12-MAR-1999) to the EMBL/GenBank/DBJ databases.
XX MIPS, Am Kioferersplitz 18a D-82152 Martinsried, GERMANY
XX
XX Clone from S. Wiemann, sequenced by Medigenomix within the CDNA
XX sequencing consortium of the German Genome Project
XX s1 sequence also available
XX This clone is available at the RZPD in Berlin
XX Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
XX Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de
XX
FH Key Location/Qualifiers
FH
FT source 1..20
FT /db_xref="taxon:9606"
FT /mol_type="mRNA"
FT /organism="Homo sapiens"
FT /clone_lib="566 (synonym: hEKd2). Vector pAMP1; host
FT X1-2b1ue; sites NotI + SalI"
FT /dev_stage="fetal"
FT /tissue_type="kidney"

```

```

XX
SQ Sequence 20 BP; 20 A; 0 C; 0 G; 0 T; 0 other;
XX
Query Match 0.7%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20
XX
RESULT 3
ID AM334823/c 20 bp mRNA linear EST 31-JAN-2000
XX LOCUS AM334823
XX DEFINITION S39F5 AGS-1 Pneumocystis carinii cDNA 3', mRNA sequence.
XX ACCESSION AM334823
XX VERSION AM334823.1 GI:6831180
XX KEYWORDS EST.
XX SOURCE
XX ORGANISM
XX Pneumocystis carinii
XX Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
XX Pneumocystidaceae; Pneumocystis.
XX 1 (bases 1 to 20)
XX Smilian, A.G., Arnold, J., Weise, M., Wunderlich, J., Staben, C.,
XX Edman, J.C., Kovacs, J. and Cushion, M.
XX Expressed sequence tags from Pneumocystis carinii
XX Unpublished (2000)
XX COMMENT
XX Contact: Staben C
XX School of Biological Sciences
XX University of Kentucky
XX 101 Morgan Building, University of Kentucky, Lexington, KY
XX 40506-0225, USA
XX Tel: 606 257 2161
XX Fax: 606 257 1717
XX Email: staben@pop.uky.edu.
XX Location/Qualifiers
XX
XX 1..20
XX /organism="Pneumocystis carinii"
XX /mol_type="mRNA"
XX /db_xref="taxon:4754"
XX /lab_host="E. coli"
XX /clone_lib="TAGS-1"
XX /note="Vector: Lambda ZAP II; Site 1: EcoRI; Site 2: XhoI;
XX P. carinii organisms (3x10e9) from a single rat (99-1-6,
XX sacrificed on 3/17/99) at Cincinnati VA facilities.
XX Trizol extracted RNA. Oligo dT priming, standard
XX conditions described by vendor. Stragene. Further
XX details see www.uky.edu/Project/Pneumocystis/"
XX
XX ORIGIN
XX
XX Query Match 0.7%; Score 20; DB 10; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 20 AAAAAAAAAAAAAAAAAA 1
XX
RESULT 4
ID CF280913 20 bp mRNA linear EST 14-AUG-2003
XX LOCUS CF280913
XX DEFINITION Oryza sativa cDNA 14ETL--07-M07, mRNA sequence.
XX ACCESSION CF280913
XX VERSION CF280913.1 GI:33658299
XX KEYWORDS EST.
XX SOURCE
XX ORGANISM
XX Oryza sativa
XX Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
XX Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
XX Ehrhartoideae; Oryzaceae; Oryza.

```

REFERENCE 1 (bases 1 to 20)
 AUTHORS Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
 TITLE Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
 JOURNAL Large-scale Sequencing Analysis of Rice ESTs
 COMMENT Unpublished (2003)
 Contact: Nahm B.H.
 Genomics and Genetics Institute, Greengene Biotech Inc.; Division
 of Bioscience and Bioinformatics, Myongji University
 Yongin, Kyonggi, Korea
 Tel: 82 31 330 6193
 Fax: 82 31 321 6355
 Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES
 source
 1..20
 /organism="Oryza sativa"
 /mol_type="mRNA"
 /cultivar="Nackdong"
 /db_xref="taxon:4530"
 /clone="14ETL--07-M07"
 /tissue_type="leaf"
 /dev_stage="14 days after germination"
 /lab_host="E.coli DH10B"
 /clone_lib="Rice etiolated leaf plasmid cDNA library
 (14ETL)"
 /note="Vector: pCR4-TOPO; Site 1: EcoRI; mRNA was capped
 with oligoribonucleotides and then used as templates for
 RT-PCR."

ORIGIN

Query Match 0.7%; Score 20; DB 14; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1799 AAAAAAAAAAAAAAAAAA 1818
 1 AAAAAAAAAAAAAAAAAA 20

RESULT 5
 LOCUS CF282035 20 bp mRNA linear EST 14-AUG-2003
 DEFINITION 14ETL--09-F19.g1 Rice etiolated leaf plasmid cDNA library (14ETL)
 ORYZA sativa cDNA clone 14ETL--09-F19, mRNA sequence.
 ACCESSION CF282035
 VERSION CF282035.1 GI:33659422
 KEYWORDS EST.
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriaraloideae; Oryzaceae; Oryza.
 1 (bases 1 to 20)
 Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
 Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
 Large-scale Sequencing Analysis of Rice ESTs
 Unpublished (2003)
 Contact: Nahm B.H.
 Genomics and Genetics Institute, Greengene Biotech Inc.; Division
 of Bioscience and Bioinformatics, Myongji University
 Yongin, Kyonggi, Korea
 Tel: 82 31 330 6193
 Fax: 82 31 321 6355
 Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES
 source
 1..20
 /organism="Oryza sativa"
 /mol_type="mRNA"
 /cultivar="Nackdong"
 /db_xref="taxon:4530"
 /clone="14ETL--09-F19"
 /tissue_type="leaf"
 /dev_stage="14 days after germination"
 /lab_host="E.coli DH10B"

/clone lib="Rice etiolated leaf plasmid cDNA library
 (14ETL)"
 /note="Vector: pCR4-TOPO; Site 1: EcoRI; mRNA was capped
 with oligoribonucleotides and then used as templates for
 RT-PCR."

ORIGIN
 Query Match 0.7%; Score 20; DB 14; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1799 AAAAAAAAAAAAAAAAAA 1818
 1 AAAAAAAAAAAAAAAAAA 20

RESULT 6
 LOCUS CF282414 20 bp mRNA linear EST 14-AUG-2003
 DEFINITION 14ETL--09-O19.g1 Rice etiolated leaf plasmid cDNA library (14ETL)
 ORYZA sativa cDNA clone 14ETL--09-O19, mRNA sequence.
 ACCESSION CF282414
 VERSION CF282414.1 GI:33659801
 KEYWORDS EST.
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriaraloideae; Oryzaceae; Oryza.
 1 (bases 1 to 20)
 Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
 Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
 Large-scale Sequencing Analysis of Rice ESTs
 Unpublished (2003)
 Contact: Nahm B.H.
 Genomics and Genetics Institute, Greengene Biotech Inc.; Division
 of Bioscience and Bioinformatics, Myongji University
 Yongin, Kyonggi, Korea
 Tel: 82 31 330 6193
 Fax: 82 31 321 6355
 Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES
 source
 1..20
 /organism="Oryza sativa"
 /mol_type="mRNA"
 /cultivar="Nackdong"
 /db_xref="taxon:4530"
 /clone="14ETL--09-O19"
 /tissue_type="leaf"
 /dev_stage="14 days after germination"
 /lab_host="E.coli DH10B"
 /clone lib="Rice etiolated leaf plasmid cDNA library
 (14ETL)"
 /note="Vector: pCR4-TOPO; Site 1: EcoRI; mRNA was capped
 with oligoribonucleotides and then used as templates for
 RT-PCR."

ORIGIN

Query Match 0.7%; Score 20; DB 14; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1799 AAAAAAAAAAAAAAAAAA 1818
 1 AAAAAAAAAAAAAAAAAA 20

RESULT 7
 LOCUS CF299822 20 bp mRNA linear EST 15-AUG-2003
 DEFINITION 7LEAF--03-P22.b1 Rice leaf plasmid cDNA library II (7LEAF) Oryza
 sativa cDNA clone 7LEAF--03-P22, mRNA sequence.
 ACCESSION CF299822

```

VERSION      CF299922.1  GI:33671583
KEYWORDS     EST.
SOURCE       Oryza sativa
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
REFERENCE    1 (bases 1 to 20)
              Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
              Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
              Large-scale Sequencing Analysis of Rice ESTs
              Unpublished (2003)
TITLE        JOURNAL
COMMENT      Contact: Nahm B.H.
              Genomics and Genetics Institute, Greengene Biotech Inc., Division
              of Bioscience and Bioinformatics, Myongji University
              Yongin, Kyeonggi, Korea
              Tel: 82 31 330 6193
              Fax: 82 31 321 6355
              Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
              Location/Qualifiers
                1..20
FEATURES     source
              /organism="Oryza sativa"
              /mol_type="mRNA"
              /cultivar="Nackdong"
              /db_xref="taxon:4530"
              /clone="7LEAF-03-P22"
              /tissue_type="leaf"
              /dev_stage="7 days after germination"
              /lab_host="E.coli DH10B"
              /clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
              /note="Vector: PCR4-TOPO; Site 1: EcoRI; mRNA was capped
              with oligoribonucleotides and then used as templates for
              RT-PCR."
ORIGIN
Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
      |||||
      20 AAAAAAAAAAAAAAAAAA 1

Db

RESULT 8
CF301720/c      20 bp mRNA linear EST 15-AUG-2003
LOCUS           7LEAF--06-L01.b1 Rice leaf plasmid cDNA library II (7LEAF) Oryza
DEFINITION      sativa cDNA clone 7LEAF--06-L01, mRNA sequence.
ACCESSION       CF301720
VERSION         CF301720.1  GI:33673481
KEYWORDS        EST.
SOURCE          Oryza sativa
ORGANISM        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
REFERENCE    1 (bases 1 to 20)
              Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
              Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
              Large-scale Sequencing Analysis of Rice ESTs
              Unpublished (2003)
TITLE        JOURNAL
COMMENT      Contact: Nahm B.H.
              Genomics and Genetics Institute, Greengene Biotech Inc., Division
              of Bioscience and Bioinformatics, Myongji University
              Yongin, Kyeonggi, Korea
              Tel: 82 31 330 6193
              Fax: 82 31 321 6355
              Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
              Location/Qualifiers
                1..20
FEATURES     source
              /organism="Oryza sativa"
              /mol_type="mRNA"
              /cultivar="Nackdong"
              /db_xref="taxon:4530"
              /clone="7LEAF-07-C16"
              /tissue_type="leaf"
              /dev_stage="7 days after germination"
              /lab_host="E.coli DH10B"
              /clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
              /note="Vector: PCR4-TOPO; Site 1: EcoRI; mRNA was capped
              with oligoribonucleotides and then used as templates for
              RT-PCR."
ORIGIN
Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
      |||||
      20 AAAAAAAAAAAAAAAAAA 1

Db

RESULT 9
CF302027/c      20 bp mRNA linear EST 15-AUG-2003
LOCUS           7LEAF--07-C16.b1 Rice leaf plasmid cDNA library II (7LEAF) Oryza
DEFINITION      sativa cDNA clone 7LEAF--07-C16, mRNA sequence.
ACCESSION       CF302027
VERSION         CF302027.1  GI:33673788
KEYWORDS        EST.
SOURCE          Oryza sativa
ORGANISM        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
REFERENCE    1 (bases 1 to 20)
              Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
              Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
              Large-scale Sequencing Analysis of Rice ESTs
              Unpublished (2003)
TITLE        JOURNAL
COMMENT      Contact: Nahm B.H.
              Genomics and Genetics Institute, Greengene Biotech Inc., Division
              of Bioscience and Bioinformatics, Myongji University
              Yongin, Kyeonggi, Korea
              Tel: 82 31 330 6193
              Fax: 82 31 321 6355
              Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
              Location/Qualifiers
                1..20
FEATURES     source
              /organism="Oryza sativa"
              /mol_type="mRNA"
              /cultivar="Nackdong"
              /db_xref="taxon:4530"
              /clone="7LEAF-07-C16"
              /tissue_type="leaf"
              /dev_stage="7 days after germination"
              /lab_host="E.coli DH10B"
              /clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
              /note="Vector: PCR4-TOPO; Site 1: EcoRI; mRNA was capped
              with oligoribonucleotides and then used as templates for
              RT-PCR."
ORIGIN
Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
      |||||
      20 AAAAAAAAAAAAAAAAAA 1

Db

RESULT 10
CF310604

```

```

              /cultivar="Nackdong"
              /db_xref="taxon:4530"
              /clone="7LEAF--06-L01"
              /tissue_type="leaf"
              /dev_stage="7 days after germination"
              /lab_host="E.coli DH10B"
              /clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
              /note="Vector: PCR4-TOPO; Site 1: EcoRI; mRNA was capped
              with oligoribonucleotides and then used as templates for
              RT-PCR."
ORIGIN
Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
      |||||
      20 AAAAAAAAAAAAAAAAAA 1

Db

RESULT 10
CF310604

```

LOCUS CF310604 20 bp mRNA linear EST 15-AUG-2003
 DEFINITION ABF--05-F14.g1 ABF3-overexpressing transgenic rice plasmid cDNA
 library (ABF) Oryza sativa cDNA clone ABF--05-F14, mRNA sequence.
 ACCESSION CF310604
 VERSION CF310604.1 GI:33682365
 KEYWORDS EST.
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 20)
 Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
 Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
 Large-scale Sequencing Analysis of Rice ESTs
 Unpublished (2003)
 COMMENT Contact: Nahm B.H.
 Genomics and Genetics Institute, Greengene Biotech Inc., Division
 of Bioscience and Bioinformatics, Myongji University
 Yongin, Kyeonggi, Korea
 Tel: 82 31 321 6355
 Fax: 82 31 321 6355
 Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES
 source
 1..20
 /organism="Oryza sativa"
 /mol_type="mRNA"
 /cultivar="Nackdong"
 /db_xref="taxon:4530"
 /clone="ABF--05-F14"
 /tissue_type="leaf"
 /dev_stage="14 days after germination"
 /lab_host="E.coli DH10B"
 /clone_lib="ABF3-overexpressing transgenic rice plasmid
 cDNA library (ABF)"
 /note="Vector: PCR4-TOPO; Site 1: EcoRI; Leaf was dried
 for 2hrs. Oligo-capped mRNA was reverse transcribed and
 then used for PCR. mRNA was prepared from ABA-responsive
 element binding transcription factor 3 overexpression
 line."

ORIGIN
 Query Match 0.7%; Score 20; DB 14; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
 |||||
 1 AAAAAAAAAAAAAAAAAA 20

RESULT 11
 LOCUS CF313067
 DEFINITION CP313067 20 bp mRNA linear EST 15-AUG-2003
 HD--01-B02.g1 OSHDAC1-overexpressing transgenic rice plasmid cDNA
 library (HD) Oryza sativa cDNA clone HD--01-B02, mRNA sequence.
 ACCESSION CF313067
 VERSION CF313067.1 GI:33684828
 KEYWORDS EST.
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 20)
 Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
 Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
 Large-scale Sequencing Analysis of Rice ESTs
 Unpublished (2003)
 COMMENT Contact: Nahm B.H.
 Genomics and Genetics Institute, Greengene Biotech Inc., Division
 of Bioscience and Bioinformatics, Myongji University
 Yongin, Kyeonggi, Korea
 Tel: 82 31 321 6355
 Fax: 82 31 321 6355
 Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES
 source
 1..20
 /organism="Oryza sativa"
 /mol_type="mRNA"
 /cultivar="Nackdong"
 /db_xref="taxon:4530"
 /clone="HD--01-B02"
 /tissue_type="callus"
 /dev_stage="proliferated callus on 2N6 media for 2 weeks"
 /lab_host="E.coli DH10B"
 /clone_lib="OSHDAC1-overexpressing transgenic rice plasmid
 cDNA library (HD)"
 /note="Vector: PCR4-TOPO; Site 1: EcoRI; Callus was
 treated with ABA(20um) for 1hr. Oligo-capped mRNA was
 reverse transcribed and then used for PCR. mRNA was
 derived from rice Histone Deacetylase overexpression
 line."

ORIGIN
 Query Match 0.7%; Score 20; DB 14; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
 |||||
 1 AAAAAAAAAAAAAAAAAA 20

RESULT 12
 LOCUS CF313569/c
 DEFINITION CF313569 20 bp mRNA linear EST 15-AUG-2003
 HD--01-L22.b1 OSHDAC1-overexpressing transgenic rice plasmid cDNA
 library (HD) Oryza sativa cDNA clone HD--01-L22, mRNA sequence.
 ACCESSION CF313569
 VERSION CF313569.1 GI:33685330
 KEYWORDS EST.
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 20)
 Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
 Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
 Large-scale Sequencing Analysis of Rice ESTs
 Unpublished (2003)
 COMMENT Contact: Nahm B.H.
 Genomics and Genetics Institute, Greengene Biotech Inc., Division
 of Bioscience and Bioinformatics, Myongji University
 Yongin, Kyeonggi, Korea
 Tel: 82 31 330 6193
 Fax: 82 31 321 6355
 Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES
 source
 1..20
 /organism="Oryza sativa"
 /mol_type="mRNA"
 /cultivar="Nackdong"
 /db_xref="taxon:4530"
 /clone="HD--01-L22"
 /tissue_type="callus"
 /dev_stage="proliferated callus on 2N6 media for 2 weeks"
 /lab_host="E.coli DH10B"
 /clone_lib="OSHDAC1-overexpressing transgenic rice plasmid
 cDNA library (HD)"
 /note="Vector: PCR4-TOPO; Site 1: EcoRI; Callus was
 treated with ABA(20um) for 1hr. Oligo-capped mRNA was
 reverse transcribed and then used for PCR. mRNA was
 derived from rice Histone Deacetylase overexpression
 line."

```

ORIGIN
Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1799 AAAAAAAAAAAAAAAAAAAAAA 1818
Db      20 AAAAAAAAAAAAAAAAAAAAAA 1

RESULT 13
LOCUS      CF319133/c      20 bp      mRNA      linear      EST 15-AUG-2003
DEFINITION HD--09-113.b1 OSHDAC1-overexpressing transgenic rice plasmid cDNA
ACCESSION  CF319133
VERSION     CF319133.1 GI:33690894
KEYWORDS   EST.
ORGANISM   Oryza sativa
SOURCE     Oryza sativa
COMMENT    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharctoidae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 20)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE     Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT    Contact: Nahm B.H.
Genomics and Genetics Institute, Greengene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
Location/Qualifiers
1..20
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="HD-09-113"
/tissue_type="callus"
/dev_stage="proliferated callus on 2N6 media for 2 weeks"
/lab_host="E.coli DH10B"
/clone_lib="OSHDAC1-overexpressing transgenic rice plasmid
cDNA library (HD)"
/notes="Vector: PCR4-TOP0; Site_1: EcoRI; Callus was
treated with ABA(20um) for 1hr. Oligo-capped mRNA was
reverse transcribed and then used for PCR. mRNA was
derived from rice Histone Deacetylase overexpression
line."

FEATURES
source

```

```

ORGANISM   Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharctoidae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 20)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE     Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT    Contact: Nahm B.H.
Genomics and Genetics Institute, Greengene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
Location/Qualifiers
1..20
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="HD-13-B05"
/tissue_type="callus"
/dev_stage="proliferated callus on 2N6 media for 2 weeks"
/lab_host="E.coli DH10B"
/clone_lib="OSHDAC1-overexpressing transgenic rice plasmid
cDNA library (HD)"
/notes="Vector: PCR4-TOP0; Site_1: EcoRI; Callus was
treated with ABA(20um) for 1hr. Oligo-capped mRNA was
reverse transcribed and then used for PCR. mRNA was
derived from rice Histone Deacetylase overexpression
line."

ORIGIN
Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1799 AAAAAAAAAAAAAAAAAAAAAA 1818
Db      20 AAAAAAAAAAAAAAAAAAAAAA 1

RESULT 15
LOCUS      CF328565      20 bp      mRNA      linear      EST 18-AUG-2003
DEFINITION NACL--03-114.b1 Rice callus plasmid cDNA library (NACL) Oryza
sativa cDNA clone NACL--03-114, mRNA sequence.
ACCESSION  CF328565
VERSION     CF328565.1 GI:33805376
KEYWORDS   EST.
ORGANISM   Oryza sativa
SOURCE     Oryza sativa
COMMENT    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharctoidae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 20)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE     Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT    Contact: Nahm B.H.
Genomics and Genetics Institute, Greengene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
Location/Qualifiers
1..20
/organism="Oryza sativa"
/mol_type="mRNA"

```

```

/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="NACL-03-114"
/tissue_type="callus"
/dev_stage="proliferated callus on 2N6 media for 30 days"
/lab_host="E.coli DH10B"
/clone_lib="Rice callus plasmid cDNA library (NACL)"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

```

ORIGIN

```

Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1799 AAAAAAAAAAAAAAAAAA 1818
      |||||
Db 1 AAAAAAAAAAAAAAAAAA 20

```

Search completed: March 7, 2004, 20:07:28
 Job time : 7092 secs